Biotechnological Advances in the Diagnosis and Treatment of Psychiatric Disorders

Canada

88-00531/

BIOTECHNOLOGICAL ADVANCES IN THE DIAGNOSIS AND TREATMENT OF PSYCHIATRIC DISORDERS

The proceedings of a national symposium sponsored by
The Mental Health Division
of Health and Welfare Canada
and
The Consultative Committee
on Mental Health Research

February 5-6-7, 1985
Toronto, Ontario, Canada

Published by the authority of the
Minister of National Health and Welfare

Opinions expressed in these articles are the responsibility of their respective authors, and do not necessarily reflect the official policy of the Department of National Health and Welfare.

Également disponible en français sous le titre "Progrès biotechnologiques dans le diagnostic et le traitement des troubles mentaux"

© Minister of Supply and Services Canada 1987

Cat. No. H34-30/1987E

ISBN 0-662-15337-5

FOREWORD AND ACKNOWLEDGEMENTS

Recent years have seen rapid advances in the biological sciences and in biotechnology, with important applications and promises for those engaged in mental health care. The goal of this symposium was to review their impact on diagnosis, treatment and service development.

Thanks are due to many people for the work and enthusiasm they brought to developing this symposium - to the Consultative Group on Mental Health Research chaired by Dr. Roger Bland of Edmonton, who initiated and developed the concept and theme, to the presenters, the session chairpersons and the planning committee, as well as to the provincial Directors of Mental Health who supported and attended the event.

The event was organized and co-ordinated by Ivy Williams of the Mental Health Division, who also edited and prepared the proceedings.

B. Wattie, Ph.D.
Director, Mental Health Division
Health and Welfare Canada

PLANNING COMMITTEE

Dr. Roger Bland, Chairman
Dr. Morton Beiser
Dr. Alan Boulton
Dr. J. Clayton
Dr. M. Dongier
Dr. D. Streiner
Dr. V. Vikis-Freibergs
Dr. Brenda Wattie
Ms. Ivy Williams

SESSION CHAIRPERSONS

Day 1

Dr. V. Vikis-Freibergs
Dr. A. Munro

Day 2

Dr. A. Boulton
Dr. G. Heseltine

Day 3

Dr. A. Leighton

The papers are grouped into three sections following the three major aims of the symposium.

The FIRST part focuses on biotechnological advances in diagnosis, treatment and clinical management.

The SECOND part examines new methodologies in drug research and development and also reviews research resources, human and financial.

In the THIRD part, some social, economic and educational aspects of research are reviewed.

While Canadian research advances rapidly in pace and complexity, the ability to communicate and digest its implications is more limited. Throughout this symposium this difficulty is addressed, and some solutions that have implications for researchers, governments and universities are suggested.

TABLE OF CONTENTS

Page

PART I. A CRITICAL REVIEW OF DIAGNOSTIC AND THERAPEUTIC DEVELOPMENTS

The Schizophrenic Disorders

Recent Advances in the Treatment of Schizophrenia and Tardive Dyskinesia, Including the Use of Neuroleptic Radioreceptor Assay
G. Chouinard, M.D. 1

Technology and the Diagnosis of Schizophrenia: The Search for a Biological Marker
M.V. Seeman, M.D. 23

The Affective Disorders

Biotechnological Advances in the Treatment of Affective Disorders
B.D. Jones, M.D. 41

The Dementias

The Clinical Diagnosis of Dementia
H. Merskey, M.D. 63

Recent Advances in the Pathological Diagnosis of Senile Dementia Alzheimer Type
M.J. Ball, M.D. 77

Alcoholism

Diagnosis and Treatment of Alcoholism
C.S. Mellor, M.D. 85

Familial Alcoholism and the Identification of High Risk Subjects
M. Dongier, M.D. 95

Page

The Anxiety Disorders

A Critical Review of Diagnostic and Therapeutic Developments: Diagnosis and Psychological Therapy
R.C. Bowen, M.D. 102

Reflections
Comments by Alastair Munro, M.D. 113

PART II. NEW BIOTECHNICAL DEVELOPMENTS

The Impact of Molecular Biology

Molecular Biology Approach to Study of the Brain Function
M. Chretien, M.D.
C. Lazure, Ph.D.
N.G. Seidah, Ph.D. 119

Radio-Immune Assay

Advances in the Development of Radioimmunoassays towards Drugs Used in the Treatment of Mental Illness
G. McKay, Ph.D.
K.K. Midha, Ph.D. 147

Electro-Physiological Diagnostic Techniques

The Convergence of Neuropsychological and EEG Measures of Cerebral Disorganization in the Endogenous Psychoses
P. Flor-Henry, M.D. 166

Molecular Biology, Cell Biology, Drug Research and Human Behaviour
J. Fraser Mustard, M.D. 174

Research and Development Issues in Canada
A.J. Moriarity, M.D. 183

Discussion of Part II
J.M. Cleghorn, M.D. 189

Page

PART III. SOCIO-ECONOMIC IMPLICATIONS AND HUMAN ASPECTS

Economic Evaluations of the Treatment of Mental Illness
Pran Manga, Ph.D. 195

Biotechnological Advances and the Design of Mental Health Systems
Michael Dear, Ph.D. 209

Teaching Critical Appraisal Skills
David L. Streiner, Ph.D. 221

Discussion and Synthesis
J.M. Cleghorn, M.D. 241

PART I. A CRITICAL REVIEW OF DIAGNOSTIC AND THERAPEUTIC DEVELOPMENTS

THE SCHIZOPHRENIC DISORDERS

RECENT ADVANCES IN THE TREATMENT OF SCHIZOPHRENIA AND TARDIVE DYSKINESIA, INCLUDING THE USE OF NEUROLEPTIC RADIORECEPTOR ASSAY

GUY CHOUINARD, M.D., M.Sc. (Pharmacol), F.R.C.P.(C), F.A.P.A.

Clinical Psychopharmacology Unit, Allan Memorial Institute,
Royal Victoria Hospital, Department of Psychiatry,
McGill University and Psychiatric Research Center,
Louis-H. Lafontaine Hospital, Montreal, Quebec

ABSTRACT

Neuroleptic therapy remains of primary importance in the inpatient and outpatient treatment of schizophrenia, not only to control psychosis and prevent relapse and readmission, but also to allow for other forms of therapies to be effective in improving the patient's social adjustment. Thus, it is hoped that the investigation of more specific dopamine receptor blocking agents and substances, which have a potentiating effect on neuroleptics, will lead to an effective treatment of schizophrenia with minimal risk of debilitating side effects. Meanwhile palliative approaches using concomitant medications may be necessary to prevent long-term complications associated with neuroleptic therapy.

PART I: RECENT ADVANCES IN THE TREATMENT OF SCHIZOPHRENIA

Schizophrenia continues to be an extremely prevalent and serious health problem. Every individual has a 1% chance of developing schizophrenia in a lifetime. In Canada, there are thirty to forty thousand episodes of schizophrenic psychosis necessitating treatment annually.

Neuroleptics offer the most efficacious treatment for both inpatient control of acute psychosis and outpatient maintenance of remission in schizophrenic individuals. The use of these drugs also increases the efficacy of other forms of treatment such as sociotherapy and family therapy. Despite the use of neuroleptic drugs however, 50% of hospital beds are occupied by patients with schizophrenia, and upon discharge, only 15% of patients remain out of hospital after two years. Furthermore, the morbidity associated with the use of neuroleptics in the form of irreversible side effects such as tardive dyskinesia is of great concern. Recent studies suggest that at least 30% of schizophrenic patients treated with neuroleptics will develop tardive dyskinesia, a potentially irreversible neurological disorder. Accurate determination of the incidence of tardive dyskinesia is hindered by the fact that neuroleptics themselves tend to mask the presence of the disorder. Furthermore, there is evidence to suggest that neuroleptics can cause dopamine receptor supersensitivity in

those regions of the brain (mesolimbic) where they have their therapeutic effect, and that this can lead to a form of rebound or supersensitivity psychosis which, in those affected, may result in a need to take the medication for the rest of their lives. These considerations have intensified the need to search for alternative forms of drug treatment for schizophrenia.

BW234U

BW234U a dimethylpiperazinyl-propylcarbozole derivative, is a potential new antipsychotic drug. Animal models (1) suggest that BW234U antagonizes post-synaptic mesolimbic dopaminergic receptors without affecting the nigrostriatal dopaminergic system. BW234U failed to elevate plasma prolactin levels over a wide dose range in humans and patients (2), indicating that dopamine (DA) receptor blockade does not occur in the hypothalamic-pituitary system. Receptor binding studies in vitro using labelled spiroperidol or DA as ligands failed to produce evidence of potent DA receptor binding properties for BW234U.

We carried out an early Phase II clinical trial of BW234U in the treatment of acute schizophrenia in newly admitted patients (2). Sixteen schizophrenic patients, newly admitted to hospital from the emergency room, underwent a three to six-day washout before being treated for four weeks with BW234U. Eight patients showed moderate to marked improvement in schizophrenic symptoms, and there was a statistically significant (p .001) reduction in the mean score for Clinical Global Impressions in all patients. Four patients did not complete the trial, two because of poor therapeutic effect, one because of a grand-mal seizure and one because of an episode of loss of consciousness of unknown origin. BW234U showed no evidence of neuroleptic plasma activity (as measured by radioreceptor assay), did not induce plasma prolactin elevation, and did not appear to cause parkinsonism. There tended to be an increase (p .10) in the total score for dyskinetic movements in patients which can be explained by the 'uncovering' effect of neuroleptic withdrawal. Also, the drug appears to have little or no sedative effect. A disadvantage is that at high doses, such as 500 mg/day, BW234U appears to affect the central nervous system, inducing tremors, muscle fasciculations, EEG abnormalities and grand-mal seizures. Controlled clinical trials of BW234U in the treatment of schizophrenia appear to be warranted.

Remoxipride

Classical neuroleptics are considered to be nonspecific for any dopaminergic region of the brain: they increase prolactin, cause parkinsonism, and have anti-Tourette, antidepressant, antimanic and antischizophrenic properties. Relative in vitro affinities for D_1 and D_2 dopamine receptors differentiate classical neuroleptics such as fluphenazine (relatively equivalent in blocking D_1 and D_2 receptors) from haloperidol which is considered to have 200 times more affinity for D_2 than D_1, and pimozide with 900 times more affinity for D_2 and D_1 (3). It is interesting to note

that remoxipride, a S(-)-3-bromo-N [1-ethyl-2-pyrrolidinyl) methyl]-2,6-dimethoxy-benzamide monohydrate, has no affinity for D_1 receptors, and thus appears to be a pure D_2 receptor blocking agent (4). Furthermore, remoxipride was found to have a preferential effect in mesolimbic and extrastriatal dopamine containing neurons (4).

We carried out an early Phase II clinical trial of remoxipride in the treatment of 20 hospitalized schizophrenic patients with and without acute exacerbation.

Twenty hospitalized schizophrenic patients, satisfying DSM-III criteria, underwent a four-day placebo washout before being treated for six weeks with remoxipride. All patients completed the clinical trial. Ten patients showed marked improvement in schizophrenic symptoms, and there was a statistically significant ($p<.05$) reduction in the mean score for CGI severity of illness and BPRS total score in all patients. Remoxipride was found to cause less parkinsonism than the prior neuroleptic therapy and appeared to have little masking effect on tardive dyskinesia. The drug showed no evidence of neuroleptic plasma activity as measured by radioreceptor assay using [^{3}H] spiperone as a ligand, did not induce plasma prolactic elevation in steady state concentrations but produced a short-lasting peak upon acute administration.

Adverse experiences reported or observed were mild in nature and included: akathisia (n=9), fatigue (n=7), insomnia (n=7). Remoxipride produced a statistically significant ($p<.05$, paired t-test) lowering of supine systolic blood pressure at the end of the clinical trial and a statistically significant ($p<.001$, paired t-test) but small decrease in mean body weight. Remoxipride appears to have little sedative effect.

Prolactin levels obtained during steady state were not very different from the mean pretreatment levels. On the other hand, there was a marked increase in prolactin in both the male and female patients one hour after drug intake confirming that remoxipride has some effect in the pituitary system.

The lack of serum neuroleptic activity measured by radioreceptor assay using [^{3}H] spiperone binding suggested that other ligands should be used for D_2 receptors in human studies of remoxipride.

In view of the results of this study which underline the importance of D_2 receptor blockage in order to achieve an antischizophrenic effect, controlled clinical trials of remoxipride in the treatment of schizophrenia also appear to be warranted.

Rubidium

Rubidium falls into group IA of the periodic table, which includes three elements of major physiologic importance (H, Na, K) and one (Li) of proven therapeutic value in manic depressive illness. Physiologically as well as chemically, Na and Li show some similarities, while K, Rb and Cs

tend to resemble each other (5). Rubidium (Rb) is present in virtually all animal tissue in the range of from 2 x 10^{-3} to 6 x 10^{-3} percent of dry weight. The concentration of K is usually some 300-800 times greater (5). The ability of Rb to substitute for K has been documented in several physiological systems (6-10). In laboratory animals deprived of dietary K, Rb can act as a temporary substitute (5). Studies show that injected or ingested Rb is readily absorbed and distributed in all the body tissues (11). Studies in man of naturally occurring Rb show that the ratio Rb/K urinary output is almost constant. Rb excretions vary from day to day, but are well correlated with long-term levels of K ingestion (11). There is thus good reason to postulate that the body tends to treat Rb as it does K, and that the mechanisms normally operative in K metabolism deal also with rubidium.

The behavioural effects of Rb are in many respects opposite to that of lithium. Lithium competes with Na, and it is reasonable to postulate that in the patient on lithium therapy there are alterations in the transmembranal sodium flux. This results in change in the electrical characteristics of the cell. A similar hypothesis may be advanced in respect of Rb. Rb can substitute for, and tends to compete with K, displacing a portion of it from its predominantly intracellular location. Intracellular-extracellular K flux plays a significant part in the electrogenesis of the nerve impulse and by substituting for K, Rb may significantly alter the properties of the neuron in a manner analogous to that of lithium. One may infer that changes in the mechanism (i.e. turnover) of transmitter substances will follow upon changes in the electrical profile of the neuron. There is some evidence to support these hypotheses based on EEG studies, studies of isolated nerve and nerve muscle preparations, and studies of individual neurons (12-14).

Rubidium in psychiatry - Rubidium has been shown to produce marked behavioural effects in animals - mainly increased activity and increased aggression. One could speculate that it could benefit withdrawn and apathetic schizophrenic patients. Its effects are in general opposite to those of lithium. Lithium does not appear to have antipsychotic properties in the majority of schizophrenic patients. Rubidium which has different physiological properties from lithium, has the potential to be beneficial for schizophrenic patients, especially for those with significant residual symptoms, and also for a great proportion of schizophrenic patients since most of them never completely go back to a normal life. Since rubidium shares with the neuroleptics their properties to increase the turnover of catecholamines (15,16), we propose that it may have a therapeutic effect on schizophrenic patients either by potentiation or as an additive effect. Up to now the three clinical trials that we have completed are compatible with this hypothesis. Neuroleptics are known to have a double effect on the metabolism of catecholamines: they block dopaminergic receptors and noradrenergic receptors, and they increase the turnover of catecholamines (17,18). However, this effect on the turnover of catecholamines of the neuroleptics seems to disappear over time, thus one could speculate that rubidium might prevent this occurring and thus have the potential to prevent the appearance of tardive dyskinesia. Another hypothesis to explain the effect of rubidium in schizophrenia could be its effect on

serotonin. Donaldson et al. found a modest but significant increase of serotonin after injection of rubidium chloride in mice (19). In fact, the role of serotonin in schizophrenia has never been completely ruled out since the finding of Wyatt et al. of an improvement with 5-hydroxytryptophan and a dopa decarboxylate inhibitor (20). Whatever its mechanism of action, it appears that rubidium needs to be further investigated, especially in view of the significant long-term side effect of neuroleptics and also due to their limited therapeutic effect over time. As regards its effect on depression, no conclusion can be reached at this time.

Fieve et al. first initiated studies of rubidium in psychiatric disorders and gave rubidium to depressed patients (21). In contrast, we postulated a potential therapeutic effect of rubidium in schizophrenia and have now carried out three double-blind controlled studies of rubidium in chronic, hospitalized schizophrenic patients, and a fourth is ongoing.

First study - This was a single-dose placebo-controlled study in which rubidium was given at four different dose levels to schizophrenic patients (Chouinard, G., Annable, L., Commun Psychopharmacol, 1977). The patients continued to take their neuroleptic maintenance medication. The purpose of this study was to determine whether rubidium had any effect and, if so, to determine the optimum dosage and the frequency that the medication should be given. The half-life of rubidium was also estimated. The results were encouraging and suggested that rubidium could increase activity and alertness in these patients and its maximum effect appeared around two to three weeks after ingestion.

Second study - We then conducted a double-blind controlled study of three months duration comparing rubidium 1.0g monthly versus rubidium placebo in schizophrenic patients receiving neuroleptic maintenance medication (Chouinard, G., New Research Abstract, APA 1979). The initial dose of rubidium appeared to cause temporary aggravation of thinking disturbance, but this disappeared following the second and third doses, and rubidium-treated patients then tended to show significant improvement relative to the control group. These findings were supported by significant correlation coefficients between rubidium serum concentrations and reduction in emotional withdrawal after two weeks, and between rubidium serum concentrations and improvement in thinking disturbance after eight weeks. To explain this biphasic response we hypothesized that rubidium gradually desensitized the dopamine-excited supersensitive neurons responsible for positive symptoms, and that the lapse in time to produce this desensitization was responsible for the delay in improvement. Rubidium appeared also to have a beneficial effect in tardive dyskinesia. Rubidium also produced some increase in prolactin, which appeared to be specific in view of the fact that its magnitude was correlated with the serum concentration of rubidium.

Third study - The third study was similar to the latter, but was of six months duration with 25% reduction in neuroleptic dosage after three months (Chouinard, G., Annable, L., Abstract, Society of Biological Psychiatry, 1983). Rubidium produced significant improvement in thinking disturbance and overall psychopathology after three months of treatment,

which persisted during the period of partial withdrawal of neuroleptics. Unlike the control group, in which neuroleptic dosage reduction resulted in a tendency to relapse requiring PRN's of chlorpromazine, rubidium-treated patients showed little signs of relapse. A biphasic treatment response with initial exacerbation and subsequent improvement of psychopathology was again observed. This supports our hypothesis explaining the mechanism of action of rubidium, that by increasing catecholamine turnover without blocking the dopamine receptor, rubidium can gradually desensitize the dopamine-excited supersensitive neurons responsible for positive schizophrenic symptoms. No evidence of toxicity was seen in any of these studies. In contrast to other treatments, such as alpha-methyldopa and alpha-methylparatyrosine, that in the past have been reported to potentiate neuroleptic effect, rubidium did not exacerbate tardive dyskinesia and, in some cases, improvement was seen. These findings are compatible with the hypothesized effect of rubidium on catecholamine turnover. No significant adverse reactions were seen in any patients. Thus, rubidium appears to be unique in that it may offer the hope of decreasing neuroleptic intake in chronic schizophrenic patients with little risk of relapse and without increasing the incidence of tardive dyskinesia.

Fourth study - We initiated a fourth study of rubidium in which a staged withdrawal of 50% of neuroleptic dosage would be made over a nine-month period in a group of chronic, hospitalized schizophrenic patients. The 32 patients have now been included, and the last group of patients will be completing the study soon. This study will allow us to assess to what extent neuroleptic medication may be decreased when rubidium is given as an adjunct.

Conclusion - We have described three drug treatment strategies which may lead to the improved treatment of schizophrenic patients. The first new drug treatment approach proposed was the use of a substance which decreased dopaminergic function in mesolimbic region of the brain by a mechanism of action differing from that of neuroleptics, meaning without dopamine receptor blockade. This approach could be the most successful since it would not only prevent the development of dopaminergic supersensitivity responsible for tardive dyskinesia in neostriatal region but would also avoid mesolimbic supersensitivity meaning supersensitivity psychosis. The second approach is one that would be successful in treating schizophrenia without the disturbing side effects associated with neuroleptic treatment of tardive dyskinesia. The drugs used in this approach share with neuroleptics the classical mechanism of dopamine blockade. However, they are more specific for D_2 dopamine receptors and for extrastriatal dopamine regions thus they would not carry the disadvantage of tardive dyskinesia but they may be associated with tardive or supersensitivity psychosis. The third approach is a palliative approach using the salt of rubidium, a trace metal, to prevent tardive dyskinesia and supersensitivity psychosis by decreasing the dose of neuroleptic necessary to produce a therapeutic effect and side effects. The advantage of this approach is that, following stabilization, patients could be given rubidium by mouth once a month and the dose of neuroleptic could be decreased by up to 50%. At this stage rubidium can produce a significant therapeutic effect with neuroleptic dose reduction of up to 25% as opposed to patients who had received rubidium

placebo. If further studies confirm the possibility of a greater diminution of neuroleptic dose this will undoubtedly constitute a major therapeutic advance preceding the introduction of other drugs which do not carry the potential of inducing long-term side effects associated with neuroleptic treatment.

PART II: TARDIVE DYSKINESIA

Neurophysiology - The most generally accepted hypothesis for the clinical manifestation of tardive dyskinesia is that of dopamine receptor supersensitivity related to prolonged blocking of dopamine postsynaptic receptors in the neostriatum (22,23). Although the dopamine system appears to interact with two other neuronal systems (cholinergic and GABA), dopamine appears to be of importance in the development and/or manifestation of tardive dyskinesia. The dopamine supersensitivity is thought to create an imbalance between the dopaminergic-cholinergic system that leads to tardive dyskinesia. Anticholinergic drugs are known to exaggerate this imbalance and increase the manifestation of the disorder and may serve as a loop for a neuronal feedback. Some authors do not agree with the dopamine supersensitivity hypothesis because of the lack of supporting evidence in man from neuroendocrine or neurophysiologic studies with CSF measurements (24). Recent knowledge about different types of dopamine receptors may help to elucidate the etiology of tardive dyskinesia.

Drugs that cause tardive dyskinesia - Although there have been reports that other drugs may cause tardive dyskinesia, the class of drugs that has most often been implicated is the neuroleptics. All types of neuroleptics (phenothiazines, butyrophenones, diphenylbutylpiperidines, thioxanthenes and benzamides (sulpiride)) have been incriminated. Metoclopramide, a selective D_2-dopamine receptor blocking agent, has also been reported to cause tardive dyskinesia (25). Rare cases have been reported with other classes of drugs (tricyclic antidepressants (26), MAO inhibitors (27), antihistamines (28), phenylhydantoin (29), antimalariales (30) and benzodiazepines (31). However, in some of these cases, the role of the presumed offending agent may only have been to uncover the disorder by its anticholinergic properties, since some of these patients had been treated previously with neuroleptics.

Another class of drugs known to induce a syndrome similar to tardive dyskinesia is the dopamine agonists. L-dopa induced dyskinesias have been well documented. Cases have also been reported with amphetamines and amantadine (32). Tardive dyskinesia and L-dopa induced dyskinesia seem to present a similar clinical picture and a common pathophysiology of an increased dopaminergic function over the cholinergic. However, L-dopa induced dyskinesias are thought to result from toxicity, are reversible and readily managed by reducing dosage and seem to involve mostly choreoathetoid movements of the extremities.

Treatment - Jeste and Wyatt recently published an extensive review of the drug trials in the treatment of tardive dyskinesia (33). We present below a brief summary of some important but neglected aspects of such

trials. We believe that drug trials in the treatment of tardive dyskinesia should be divided into two categories: short- and long-term. The aim of the short-term trial is to find a drug that acts rapidly to suppress or alleviate the syndrome and help the most handicapped patients. The long-term effects of such a treatment are unknown and also need to be studied. There is also the possibility that a drug may exacerbate the disorder in the short-term, but prevent it in the long-term. Long-term trials are also needed for agents that may prevent development of the disorder. Short-term trials have been carried out with drugs that affect either dopaminergic, cholinergic or GABA neurotransmission, but without consistent success. Clearly, the most efficacious treatment on a short-term basis has been the use of the high potency neuroleptics, especially when given on a QID regimen. However, in the long-term such treatment will of course only compound the problem. As regards cholinergic drugs, we have summarized the clinical trials of deanol and reported a lack of superiority of this drug over placebo (34). Long-term studies of deanol seem unwarranted since it is doubtful that it increases cholinergic function. Other acetylcholine (ACH) precursors, choline and lecithine, have been studied. The most promising approach seems to be the use of L-dopa which may act to desensitize supersensitive receptors. Alpert and his co-workers have recommended such an approach and recently reported encouraging findings (35). We have conducted a short-term clinical trial of L-dopa in the treatment of drug-induced extrapyramidal symptoms (36). The results suggest that L-dopa is significantly better than placebo in the treatment of parkinsonian symptoms. As expected, tardive dyskinesia was increased by the short-term administration of L-dopa. However, by the same mechanism that L-dopa is found to increase tardive dyskinesia in the short-term, it may, as suggested by Alpert and Friedhoff, be of therapeutic value in the long-term and prevent the development of the disorder. Also, the increase in tardive dyskinesia observed in our short-term trial was much less than that which occurred when the patients were subsequently switched to an anticholinergic antiparkinsonian (36). Another class of drug that has been tried is the GABA agonists such as sodium valproate, baclofen and muscinol (37-39). Results have however been equivocal. Numerous other drugs have been tried without consistent results. To summarize, no treatment has yet been found that is effective on a long-term basis.

Advantages of L-dopa - It is possible that L-dopa may present advantages over the anticholinergic agents currently utilized as antiparkinsonian drugs in the treatment of neuroleptic-induced parkinsonian symptoms. L-dopa may have a greater effect on certain symptoms induced by the neuroleptics, such as akinesia (this is true in the case of Parkinson's disease). The possibility that a more specific treatment of extrapyramidal symptoms might prevent the appearance of tardive dyskinesias which have been associated with neuroleptic administration represents another potential advantage. L-dopa could effectively prevent the development of dopaminergic receptor supersensitivity to the neurotransmitter, which is thought to be the probable cause of tardive dyskinesia. Recent animal studies suggest that dopamine agonists might prevent the development of dopaminergic supersensitivity associated with neuroleptic use (40) and that L-dopa reverses dopamine supersensitivity induced by reserpine (41). Thus, L-dopa may have the potential to prevent dopaminergic supersensitivity

induced by different classes of neuroleptics (dopamine blocking agents or depleting agents). This effect appears rather specific to neuroleptics since L-dopa does not have an effect on supersensitivity resulting from 6-hydroxydopamine striatal lesions in rats (42). Previous open clinical trials of L-dopa in the treatment of tardive dyskinesia suggest a potential therapeutic effect (43-45). It would thus appear that a placebo-controlled study is necessary.

Effectiveness in drug-induced parkinsonism - The efficacy of L-dopa in the treatment of neuroleptic-induced parkinsonian symptoms has yet to be established. To our knowledge, five studies regarding L-dopa have been published to date (46-50). The first study reports that only four patients out of 22 studied were considered as improved (46). The second study of 40 patients (20 men and 20 women) reports that the administration of a single I.V. dose of L-dopa provided a great improvement (47). The third study was not performed under controlled conditions and includes only four patients, none of whom seemed to derive any benefit from the L-dopa-benserazide combination (48). With respect to the fourth trial (49), a decrease in extrapyramidal symptoms following treatment with the L-dopa benserazide combination was reported in the four patients in whom these symptoms were present before treatment began despite administration of anticholinergic antiparkinsonian agents. The last report concerns a study which included six patients for whom L-dopa added to the neuroleptic did not improve parkinsonism (50). We carried out the first double-blind placebo- controlled study in which the effect of L-dopa on neuroleptic-induced extrapyramidal symptoms and schizophrenic symptoms was studied (36). Twenty schizophrenic inpatients with drug-induced parkinsonism were randomly assigned to treatment with either L-dopa or placebo. L-dopa was given with benserazide, a dopa carboxylase inhibitor. The three-week trial began one week after discontinuation of the patients' antiparkinsonian treatment. At the same time, an equivalent dose of fluphenazine HCl was also substituted for the patients' regular neuroleptic medication. Psychiatric (BPRS and CGI) and extrapyramidal symptoms were measured weekly. Patients treated with L-dopa had 1) a significantly ($p<.05$) lower adjusted final mean score for hyperkinetic parkinsonian symptoms than those treated with placebo, 2) a significant ($p<.05$) reduction in hypokinetic parkinsonian symptoms, and 3) a tendency for less ($p<.10$) positive schizophrenic symptoms than those treated with placebo. Following the trial, all patients received procyclidine 30 mg/day for two weeks. After the switch to procyclidine the psychopathology of patients previously treated with L-dope deteriorated, while that of patients previously treated with placebo improved. Procyclidine also increased dyskinetic movements in both groups. L-dopa was thus effective in reducing neuroleptic-induced parkinsonism without exacerbating psychosis or uncovering tardive dyskinesia to the extent that occurred subsequently with procyclidine.

L-dopa and schizophrenia - It appears that the reluctance of physicians to use L-dopa in the treatment of neuroleptic-induced parkinsonian symptoms stems from the general belief that L-dopa may exacerbate schizophrenic psychosis. Although anticholinergic drugs are an efficacious treatment of neuroleptic-induced parkinsonism, they do not prevent the development of tardive dyskinesia (51). L-dopa, another type of

antiparkinsonian drug, has been discarded following an uncontrolled study which reported that it was inefficacious and exacerbated psychotic symptoms (48). This early anecdotal report including only four patients had major weaknesses since patients had drug-induced parkinsonism unresponsive to anticholinergics; in addition abrupt withdrawal of anticholinergics was used which may have caused increased psychosis associated with exacerbation of parkinsonian symptoms (50,52). It is worth noting that in a second report on the same study, the author is more cautious in his conclusions (53). In contrast, other reports from both uncontrolled (54-56) and controlled studies (49,57) have found a therapeutic effect for L-dopa when given in association with neuroleptics in the treatment of schizophrenia. These studies included a greater number of patients (42 and 80 patients) and their design is more appropriate. In these studies worsening of psychosis was noted in 0% and 6% of patients, respectively, whereas 20% and 40% showed no change and all the other patients (80% and 54%) were considered from slightly to greatly improved. A recent report of exacerbation of psychosis is derived from an uncontrolled study of three patients, two of whom showed an increase of psychosis (58). However, on entering the trial these two patients had a severe illness that was not adequately controlled by neuroleptic medication. Under these conditions the L-dopa may have caused an exacerbation of the psychosis.

Another study reported on the effects of L-dopa alone in schizophrenia (59). This open study included 10 patients suffering from acute schizophrenia. Although seven of the 10 patients showed deterioration of their condition, it is not possible to compare this effect with that of placebo treatment as no control group was included. In contrast, a more recent study does not report exacerbation of schizophrenia with L-dopa (60). In addition, a publication concerning a Japanese study reports deterioration in 4% of schizophrenic patients treated with L-dopa, as compared to 22% of patients in whom an improvement was noted (61). Finally, we completed a double-blind controlled study in which L-dopa was, on the contrary, found to improve psychosis (36). Alpert and Freidhoff (43) suggested the use of L-dopa in the prevention of tardive dyskinesia and reported encouraging results without exacerbation of psychosis (62).

Study design - As a result of the need to assess the long-term effect of L-dopa on drug-induced extrapyramidal symptoms and schizophrenic symptoms, we carried out a 6-month placebo-controlled study of L-dopa which was followed by a 10-week withdrawal phase during which the effects of L-dopa were examined further. All patients included in the study had signs of at least borderline tardive dyskinesia in one or more body areas while treated with a standard neuroleptic (fluphenazine) and a standard anticholinergic antiparkinsonian (procyclidine). In addition, they were also required to present significant drug-induced parkinsonism while off the anticholinergic antiparkinsonian for one week. Since it was also important to assess the presence of masked dyskinesia, there was a non-blind 2-week treatment with procyclidine at the end of the 10-week withdrawal period. We think that the study population proposed was appropriate because when covert dyskinesias are taken into account, it has been estimated that 50-70% of neuroleptic-treated patients have tardive dyskinesia at some stage of development (63). Thus, we expect to be able to assess the therapeutic

effect of L-dopa more accurately by masking covert forms of dyskinesia by means of procyclidine administration. Patients who exhibited drug-induced parkinsonism reflected the fact that the medication was at a significant concentration in the brain; as there is no accurate way to measure neuroleptic levels in the brain, the inclusion of such patients provided the best option to ascertain that the medications were at a sufficient concentration in the neostriatum. It is also possible that this particular group of patients was at a higher risk to develop tardive dyskinesia. This study is of great interest since it may open up a new drug treatment strategy by the use of dopamine agonists (L-dopa and others) in the treatment of schizophrenia. Again, this approach is a palliative one, for the interim preceding the introduction of other antischizophrenic substances.

PART III: NEUROLEPTIC RADIORECEPTOR ASSAY (nRRA)

We have also conducted studies to evaluate the clinical utility of the radioreceptor assay for neuroleptics (nRRA) in monitoring the treatment of schizophrenia. These studies examined the effect of this monitoring of neuroleptic plasma levels (activity) on outpatients who were in remission and requiring maintenance therapy, and on inpatients who were acutely psychotic. The objective was to determine whether monitoring neuroleptic therapy by plasma level can reduce morbidity associated with neuroleptic treatment, lessen relapse rate and decrease length of hospitalization.

Plasma level monitoring of neuroleptic drugs would seem to be clinically indicated for the following reasons:

- Neuroleptic toxicity is readily confused with the underlying disease (akinetic depression, akinetic catatonia, akathisic agitation, akathisic insomnia).

- Potentially irreversible side effects of neuroleptics such as tardive dyskinesia (and supersensitivity psychosis) may be associated with high plasma levels.

- Compliance with medication is a particularly important problem in the patient population treated with neuroleptics.

- Neuroleptics do not demonstrate significant antipsychotic effect until two to three weeks after the drug has been started thus making it difficult to assess effective dose by observing clinical improvement.

- Neuroleptics for many patients are poorly or erratically absorbed leading to poor correlation between administered dose and plasma concentration.

Up until recently, blood assays of neuroleptics have not been applicable to clinical practice due to cost, or excessive time required to measure a sample as a result of complex analytic methods such as gas chromatography. Simpler assay methods such as radioimmunoassay avoid some

of these problems but are restricted in utility by their measurement of one specific neuroleptic, not metabolites of that neuroleptic which can be active. In contrast the radioreceptor assay for neuroleptics (nRRA) has been proposed for clinical use in part because it may be able to estimate the total of neuroleptic activity; that is, the activity contributed by both parent compound and metabolites, in blood and tissue specimens (64,65). As the nRRA is sensitive in vitro to neuroleptic agents of diverse chemical classes (66,67,68), it may detect neuroleptic activity in the blood of a patient on any neuroleptic. Thus, this one assay technique might be used to monitor blood neuroleptic levels in patients on various neuroleptics. For these reasons, this assay may be potentially applicable to clinical practice. Up to 100 samples can be run in a few hours thus expediting turn-around times so that results can be used in clinical decision making.

At present, measurements of neuroleptics using the nRRA have been shown to correlate with clinical improvement in psychotic patients using variable (69) and fixed doses (70) of neuroleptics. A minimally effective plasma level has been established for a number of neuroleptics (71). What remains to be determined is whether these research findings can be applied to clinical practice to effectively improve the in-hospital and out-of-hospital treatment of psychotic patients. Such improvement could occur in the following ways:

- Relapse after stabilization of the illness and discharge from hospital could be lessened by maintaining compliance and maintaining optimal drug levels.

- Short-term and long-term side effects and morbidity from the drugs may be lessened due to the ability to maintain minimally effective drug plasma levels for patients requiring long-term treatment.

- Acutely psychotic inpatients could require less hospitalization time if optimal plasma levels were attained early in the course of treatment.

- The risk of toxicity during hospitalization could be decreased.

These improvements are of obvious importance in reducing mental health costs and in allowing the patient to experience the benefits of drug therapy while minimizing the risks. The purpose of our studies was to evaluate the actual improvements accrued through the use of the nRRA in the outpatient and inpatient care of schizophrenic patients.

With the advent of neuroleptic medication the pattern of hospitalization has changed for schizophrenic patients. The duration of each hospital experience diminished. However the number of readmissions increased. At present 50% of all admissions to mental hospitals are readmissions. Statistics have also shown that a patient discharged from a psychiatric hospital has a 50-50 chance of staying out one year and only 15% of patients stay out for more than two years without being readmitted (72).

Neuroleptic treatment and compliance with that treatment has been shown repeatedly to be of crucial importance in controlling the acute psychotic phase of the illness in hospital (73) and in preventing relapse and readmission once out of hospital (74). However, from the above statistics cited, it is evident that despite widespread use of neuroleptics, schizophrenic patients continue to occupy more hospital beds than other psychiatric illnesses combined and that readmission after discharge from hospital remains a significant risk even in patients treated with medication.

Despite efforts to institute other forms of therapy such as sociotherapy and family therapy, neuroleptic treatment continues to be the most successful single treatment for schizophrenia and has been shown clearly to be necessary before other forms of intervention can be effective (75,76). Thus at the present time, it remains of utmost importance that neuroleptic treatment continue for most patients and that effective monitoring of that treatment be promoted.

This importance is emphasized when one considers the risks of long term, at times irreversible side effects identified in patients treated with neuroleptics. Tardive dyskinesia, an irreversible neurological side effect results in abnormal involuntary movement of different muscle groups in a patient exposed to neuroleptics (77). Factors related to tardive dyskinesia development have been identified by epidemiological studies (78). However, drug dose or history of drug exposure has not been found to correlate with the appearance of tardive dyskinesia (79). A recent study may explain in part this surprisingly negative finding. Jeste et al. (80) noted relatively high levels of neuroleptics in blood at correspondingly low or normal dose levels in patients who had developed tardive dyskinesia. Thus an idiosyncratic metabolic or absorptive response to neuroleptics may be present in some patients who develop tardive dyskinesia. This idiosyncratic response would only be detectable with the use of plasma measurements of neuroleptics to delineate those patients at risk if medication is not subsequently lowered and eventually discontinued. This idiosyncratic response might also explain why some patients do not respond to standard doses of neuroleptics. We have observed abnormally high neuroleptics levels in the early phase of the present study in some patients, which were manifested in toxic behavioural syndromes that remit upon decrease of medication. Clinically it was impossible to distinguish this toxicity from symptoms of the original illness.

Supersensitivity psychosis is another long term, possibly irreversible side effect of neuroleptic treatment recently identified by our group (81, 82, 83). Some patients on neuroleptics for long periods of time begin to manifest tolerance to the antipsychotic effect of the medication. This can lead to relapse if the dose is not increased. However, increasing the medication indiscriminately can lead to further tolerance and finally an irreversible need for medication despite the course of illness. The measurement of plasma levels of neuroleptics would allow for early detection of actual tolerance to the medication's effect that is central rather than peripheral due to decreased absorption of drug or due to lack of compliance with treatment. This detection would allow for treatment to be slowly

decreased at a stage where the drug-induced psychosis is still reversible. For those patients who do not have this side effect, the nRRA will facilitate the establishment of the minimal therapeutic dose.

The radioreceptor assay technique has arisen out of work demonstrating that all active neuroleptics have an affinity, related to their potency, for binding sites in brain that may represent the dopamine (DA) receptor (64, 65, 67, 68). Competition among neuroleptics for the neuroleptic receptor binding sites has been shown by demonstrating displacement of a labeled ligand (neuroleptic) from binding sites by unlabeled neuroleptic. Neuroleptic affinity for binding sites (i.e., ability to displace labeled neuroleptic) correlates closely with in vivo behavioural effects in animals and with milligram antipsychotic potency in patients (84). This is consistent with the theory that such binding, and possibly DA blockade, may underlie the specific antipsychotic mechanism of neuroleptics.

The major advantage of this assay system is that it should measure all active neuroleptic agents, original drug or metabolite, present within the limits of detectability in a specimen. In other words, the technique does not measure any neuroleptic specifically but, rather, measures neuroleptic activity in a specimen with active neuroleptic being determined by its ability to compete with a labeled ligand for tissue binding sites.

Clinical studies have demonstrated that a wide range of neuroleptics of different chemical classes can be successfully detected in blood specimens via the radioreceptor assay (69,85). Radioligand binding techniques allow a distinction to be made between typical and atypical neuroleptics in terms of their affinities for DA receptors thus enabling the clinician to choose drugs with low affinities for certain receptors in an attempt to minimize adverse effects resulting from receptor blockade.

In summary, it is evident that neuroleptic therapy remains of primary importance in the treatment of schizophrenia both as inpatient and outpatient not only to control psychosis and prevent relapse and readmission but also to allow for other forms of therapies to be effective in improving the patient's social adjustment. Thus, it is hoped that the investigation of more specific dopamine receptor blocking agents and substances which have a potentiating effect on neuroleptics, will lead to an effective treatment of schizophrenia with minimal risk of debilitating side effects. Meanwhile palliative approaches using concomitant medications such as rubidium and L-dopa may be necessary to prevent long term complications associated with neuroleptic therapy. Monitoring neuroleptic activity in plasma with the nRRA may also permit a better clinical use of neuroleptic drugs.

REFERENCES

1. McKenzie, G.M.: Apomorphine-induced aggression in the rat. Brain Res 34:323-330, 1971.

2. Chouinard, G., Annable, L.: An early Phase II clinical trial of BW234U in the treatment of acute schizophrenia in newly admitted patients. Psychopharmacology 84:282-284, 1984.

3. Christensen, J., Arnt, J., Hyttel, J. et al.: Pharmacological effects of a specific dopamine D-1 antagonist SCH 23390 in comparison with neuroleptics. Life Sci 34:1529-1540, 1984.

4. Ogren, S.O., Hall, H., Kohler, C., Magnusson, O., Lindbom, L-O, Angeby, K., Florvall, L.: Remoxipride, a new potential antipsychotic compound with selective anti-dopaminergic actions in the rat brain. European Journal of Pharmacology, 102:459-474, 1983.

5. Relman, A.S.: The physiological behavior of rubidium and cesium in relation to that of potassium. Yal J Biol Med 29:248-602, 1956.

6. Feng, T.P., Lui, Y.M.: The concentration-effect relationship in the depolarization of amphibian nerve by potassium and other agents. J Cell Physiol 34:33-42, 1949.

7. Tarail, R., Bennett, T.E., Noell, W.K.: A Comparison of electrocardiographic, electroencephalographic, and chemical effects of rubidium and potassium in dogs. J Clin Invest 34:966, 1955.

8. Mann, P.J.G., Tennenbaum, M., Quastel, J.H.: Acetylcholine metabolism in the central nervous system. The effects of potassium and other ions on acetylcholine liberation. Biochem J 33:822-835, 1939.

9. Huf, E.G., Wills, J.: Influence of some inorganic actions on active salt and water uptake by isolated frog skin. Am J Physiol 167:255-260, 1951.

10. Von Korff, R.W.: The effects of alkali metal ions of the acetate activating enzyme system. J Biol Chem 203:265-271, 1953.

11. Meltzer, H.L., Lieberman, K.W., Shelley, E.M., Stallone, F., Fieve, R.R.: Metabolism of naturally occurring Rb in the human: the constancy of urinary Rb/K. Biochem Med 6052, BK-422, Gal. 1-Gal. 3, 1972.

12. Barratt, E.S., Creson, D.L., Russell, G.: The effects of lithium salts on brain activity in the cat. Am J Psychiatry 125:530-536, 1968.

13. Tupin, J.P., Schlagenhauf, G.K., Creson, D.L.: Lithium effects on electrolyte excretion. Am J Psychiatry 125:536-543, 1968.

14. Guyton, A.C.: Textbook of Medical Physiology, 4th Edition, Philadelphia, London, Toronto: Saunders WB Co., p. 66-72, 1971.

15. Stolk, J., Barchas, J.D., Vernikos-Danellis, J.: Increased turnover of rat brain norephinephrine during rubidium treatment. Fed Proc 29:511, 1970.

16. Stolk, J.M., Norwack, W.J., Barchas, J.D.: Brain norepinephrine: enhanced turnover after rubidium treatment. Science 168:501-503, 1970.

17. Nyback, H., Borzecki, Z., Sedwell, G.: Accumulation and disappearance of catecholamines formed from tyrosine 14-C in mouse brain: Effect of some psychotropic drugs. Eur J Pharmacol 4:395-403, 1968.

18. Laverty, R., Sharman, D.F.: Modification by drugs of the metabolism of 3,4-dihydroxyphenylethylamine, noradrenaline and 5-hydroxytryptamine in the brain. Br J Pharmacol 24:759- 822, 1965.

19. Donaldson, J., Minnich, J.L., Barbeau, A.: Effet du lithium et du rubidium sur les monamines cerebrales de la souris. Union Med Can 102:907-911, 1973.

20. Wyatt, R.J., Vaughan, T., Galanter, M., Kaplan, J. Green, R.: Behavioral changes of chronic schizophrenic patients given L-5-hydroxytryptophan. Science 177:124-1126, 1972.

21. Fieve, R.R., Meltzer, H., Dunner, D.L., Levitt, M., Mendlewicz, J. Thomas, A.: Rubidium: biochemical, behavioral and matabolic studies in humans. Am J Psychiatry 130:55-61, 1973.

22. Carlsson, A. Biochemical implications of dopa-induced actions on the central nervous system with particular relevance to abnormal movements. In: L-dopa and parkinsonism. A. Barbeau & F.J. McDowel(eds), pp. 205-213. Philadelphia: Davis, 1970.

23. Klawans, H.L. The pharmacology of tardive dyskinesia. Am J Psychiatry 130:82-86, 1973.

24. Bowers, M.B. Jr., Moore, D., Tarsy, D. Tardive dyskinesia: A clinical test of the supersensitivity hypothesis. Psychopharmacology 61:137-141, 1969.

25. Lavy, S., Malamed, E., Penchas, S. Tardive dyskinesia associated with metoclopramide. Br Med J 1:77-78, 1978.

26. Fann, W.E., Sullivan, J.L., Richman, B.W. Dyskinesias associated with tricylcic antidepressants. Br J Psychiatry 128:490-493, 1976.

27. Stancer, H.C. Tardive dyskinesia not associated with neuroleptics. Am J Psychiatry 136:727, 1979.

28. Thach, B.T., Chase, T.N., Bosma, J.F. Oral facila dyskinesia associated with prolonged use of antihistaminic decongestants. N Engl J Med 293:486-487, 1975.

29. Chadwick, D., Reynolds, E.H., Marsden, C.D. Anticonvulsant-induced dyskinesias: A comparison with dyskinesias induced by neuroleptics. J Neurol Neurosurg Psychiatry 39:1210-1218, 1976.

30. Osifo, N.G. Drug-related transient dyskinesias. Clin Pharmacol Ther 25:767-771, 1979.

31. Kaplan, S.R., Murkofsky, C. Oral-buccal dyskinesia symptoms associated with low-dose benzodiazepine treatment. Am J Psychiatry 135:1558-1559, 1978.

32. Kazamatsuri, H., Chien, C.P., Cole, J.O. Therapeutic approaches to tardive dyskinesia. Arch Gen Psychiatry 27:491-499, 1972.

33. Jeste, D.V., Wyatt, R.J. In search of treatment for tardive dyskinesia: Review of the literature. Schizophrenia Bull 5:251-293, 1979.

34. de Montigny, C., Chouinard, G., Annable, L. Ineffectiveness of deanol in tardive dyskinesia: A placebo controlled study. Psychopharmacology 65:219-223, 1979.

35. Alpert, M., Friedhoff, A.J., Marcos, L.R., Diamond, F. Paradoxical reaction to L-dopa in schizophrenic patients. Am J Psychiatry 135:1329-1332, 1978.

36. Chouinard, G., Jones, B.D., Annable, L. L-dopa in drug-induced extrapyramidal symptoms. 35th Annual Convention & Scientific Program of the Society of Biological Psychiatry, Boston, Mass., Sept. 5-7, 1980.

37. Casey, D.E., Hammerstad, J.P. Sodium Valproate in tardive dyskinesia. J Clin Psychiatry 40:483-485, 1979.

38. Gerlach, J., Rye, T., Kristjansen, P. Effect of baclofen on tardive dyskinesia. Psychopharmacology 57:145-151, 1978.

39. Tamminga, C.A., Crayton, J.W., Chase, T.N. Improvement in tardive dyskinesia after muscimol therapy. Arch Gen Psychiatry 36:595-598, 1979.

40. Christensen, A.V., Nielsen, I.M. Dopaminergic supersensitivity: Influence of dopamine agonists, cholinergics, anticholinergics, and drugs used for the treatment of tardive dyskinesia. Psychopharmacology 62:111-116, 1979.

41. Ezrin-Waters, C., Seeman, P. L-dopa reversal of hyperdopaminergic behavior. Life Sci 22:1027-1032, 1978.

42. Suga, M. Effect of long-term L-dopa administration of the dopaminergic and cholinergic (muscarinic) receptors of striatum in 6-hydroxydopamine lesioned rats. Life Sci 27:877-882, 1980.

43. Alpert, M., Friedhoff, A.J. Receptor sensitivity modification in the treatment of tardive dyskinesia. Clin Pharmacol Ther 19:103, 1976.

44. Carroll, B.J., Curtis, G.C., Kokmen, E. Paradoxical response to dopamine agonists in tardive dyskinesia. Am J Psychiatry 134:785-789, 1977.

45. Friedhoff, A.J., Bonnet, K., Rosengarten, H. Reversal of two manifestations of dopamine receptor supersensitivity by administration of L-dopa. Res Commun Chem Pathol Pharmacol 16:411-423, 1977.

46. McGeer, P.L., Gibson, W.C., Foulkes, R.G. Drug-induced extrapyramidal reactions. J Am Med Ass 177:101-106, 1961.

47. Bruno, A., Bruno, S.C. Effects of L-dopa on pharmacological parkinsonism. Acta Psychiat Scand 42:264-271, 1966.

48. Yaryura-Tobias, J.A., Wolpert, A., Dana, L., Merlis, S. Action of L-dopa in drug-induced extrapyramidalism. Dis Nerv Syst 31:60-63, 1970.

49. Gerlach, J., Luhdorf, K. The effect of L-dopa on young patients with simple schizophrenia, treated with neuroleptic drugs. Psychopharmacology 44:105-110, 1975.

50. Fleming, P., Makar, H., Hunter, K.R. Levodopa in drug-induced extrapyramidal disorders. Lancet II:1186, 1970.

51. Chouinard, G., de Montigny, C., Annable, L. Tardive dyskinesia and antiparkinsonian medication. Am J Psychiatry 136:228:229, 1979.

52. Alpert, M., Diamond, F., Laski, E. Anticholinergic exacerbation of phenothiasine-induced extrapyramidal syndrome. Am J Psychiatry 133:1073-1075, 1976.

53. Yaryura-Tobias, J.A., Wolpert, A., Dana, L., Merlis, S. The action of L-dopa on schizophrenic patients. Curr Ther Res 12:528-531, 1970.

54. Inanaga, K., Oshima, M., Tachibana, H., Nakamura, K., Koketsu, K. Three cases of schizophrenia treated with L-dopa. Kurume Med J 18:161-168, 1971.

55. Garfinkel, P.E., Stancer, H.C. L-dopa and schizophrenia. Can Psychiatric Assoc J 21:27-29, 1976.

56. Inanaga, K., Inoue, K., Tachibana, H., Oshima, M., Kotorii, T. Effect of L-dopa in schizophrenia. Folia Psychiatr Nerol Jpn 26:145-157, 1972.

57. Inanaga, K., Nakazawa, Y., Inoue, K. et al. Double-blind controlled study of L-dopa therapy in schizophrenia. Folia Psychiatr Neurol Jpn 29:123-143, 1975.

58. Calil, H.M., Yesavage, A., Hollister, L.E. Low dose levodopa in schizophrenia. Commun Psychopharmacol 1:593-596, 1977.

59. Angrist, B., Sathananthan, G., Gershon, S. Behavioral effects of L-dopa in schizophrenic patients. Psychopharmacolgia (Berl) 31:1-12, 1973.

60. Brambilla, F., Scarone, S., Ponsano, M., Maffei, C., Nobile, P., Rovere, C., Guastalla, A. Catecholaminergic drugs in chronic schizophrenia. Neuropsychobiology 5:185-200, 1979.

61. Ogura, C., Kishomoto, A., Nakao, T. Clinical effect of L-dopa on schizophrenia. Curr Ther Res 20:308-318, 1976.

62. Alpert, M., Friedhoff, A. An un-dopamine hypothesis. Syllabus and Scientific Proceedings, 133rd Annual Meeting of the American Psychiatric Association, p. 25, 1980.

63. Carpenter, W.T., Rey, A.C., Stephens, J.H. Covert dyskinesia in ambulatory schizophrenia. Lancet II:212-213, 1980.

64. Cohen, B.M., Herschel, M., Aoba, A. Neuroleptic antimuscarinic, and antiadrenergic activity of chlorpormazine, thioridazine and their metabolites. Psychiatry Research, 1:199, 1979.

65. Creese, I., Snyder, S.H. A novel, simple and sensitive radioreceptor assay for anti-schizophrenic drugs in blood. Nature, 270, 180, 1977.

66. Baldessarini, R.J., Cohen, B.M., Herschel, M., Campbell, A., Lipinski, J.F. Radioreceptor assay for neuroleptics. Am Psychiat Assoc. New Research, Abs. NR36, 1979.

67. Burt, D.R., Creese, I., Snyder, S.H. Properties of (^{3}H) haloperidol and (^{3}H) dopamine binding associated with dopamine receptors in calf brain membranes. Molecular Pharmacology 12:800, 1976.

68. Seeman, P., Chau-Wong, M., Tedesco, J., Wong, K. Brain receptors for antipsychotic drugs and dopamine: Direct binding assays. Proceedings of the National Academy of Sciences of the United States of America, 72, 4376, 1975.

69. Cohen, B.M., Lipinski, J.F., Harris, P.Q., Pope, H.G., Jr., Friedman, M. Clinical use of the radioreceptor assay for neuroleptics. Psychiatry Research 1:173, 1980.

70. Cohen, B.M., Lipinski, J.F. Radioreceptor assay in clinical psychiatric research. In: Usdin E. ed Neuroreceptors - Basic and Clinical Aspects. John Wiley and Sons Ltd., New York, 1980.

71. Cohen, B.M. Personal Communication (1981).

72. Anthony, W.A., Buell, G.J., Sharrat, S., Althoff, M.E. Efficacy of psychiatric rehabilitation. Psychological Bulletin 456-477 (1972-1978).

73. Davis, J.M., Garver, D.L. Neuroleptics: clinical use in psychiatry. In: Handbook of Psychopharmacology, Vol. 10, Iverson, L.L., Iverson, S.D., and Snyder, S.H. (eds) Plenum Press, New York (1978), pp. 129-164.

74. Davis, J.M. Overview: Maintenance therapy in psychiatry. I Schizophrenia. Am J Psychiatry 132:1237, 1975.

75. Vaughan, C.E., Leff, L.P. The influence of family and social factors on the course of psychiatric illness: a comparison of schizophrenic and depressed neurotic patients. Bri J Psychiatry 129, 125, 1976.

76. Goldberg, S.C., Schooler, N.R., Hogarty, E.G., Roper, M. Prediction of relapse in schizophrenic outpatients treated by drug and sociotherapy. Arch Gen Psychiatry 34:171, 1977.

77. Tarsy, D., Baldessarini, R.J. The tardive dyskinesia syndrome. In: Klawans, H.L. (ed.). Clinical Neuropharmacology, Vol. 1, 29 (1976).

78. Chouinard, G., Annable, L., Ross-Chouinard, A., et al. Factors related to tardive dyskinesia. Am J Psychiatry 136:79, 1979.

79. American Psychiatric Association Task Force Report on Tardive Dyskinesia, 1980.

80. Jeste, D., Rosenblatt, J.F., Wagner, R.L., Wyatt, R.J. High serum neuroleptic levels in tardive dyskinesia. N.E.J.M. 301:1184, 1979.

81. Chouinard, G., Jones, B.D. Neuroleptic-induced supersensitivity psychosis: clinical and pharmacologic characteristics. Am J Psychiatry 137:16, 1980.

82. Chouinard, G., Jones, B.D. Defining drug-induced supersensitivity. Reply. Am J Psychiatry 137:992, 1980.

83. Chouinard, G., Jones, B.D., Annable, L. Neuroleptic-induced supersensitivity psychosis. Am J Psychiatry 135:1409-1410, 1978.

84. Creese, I., Burt, D.R., Snyder, S.H. Dopamine receptor binding predicts clinical and pharmacological potencies of antischizophrenic drugs. Science 192:481, 1976.

85. Calil, H.M., Avery, D.H., Hollister, L.E., Creese, I., Snyder, S.H. Serum levels of neuroleptics measured by dopamine radioreceptor assay and some clinical observations. Psychiatry Research 1:39, 1979.

TECHNOLOGY AND THE DIAGNOSIS OF SCHIZOPHRENIA: THE SEARCH FOR A BIOLOGICAL MARKER

MARY V. SEEMAN, M.D
Head of Psychiatry
Mount Sinai Hospital
Toronto

ABSTRACT

Four hypotheses of schizophrenia--the dopamine, norepinephrine, viral, and regional--are examined with respect to new findings resulting from new technological advances.

Markers of dopamine and norepinephrine activity, brain atrophy, relative hypofrontality, smooth eye pursuit dysfunction and relative left hemisphere impairment all seem promising but are, as yet, not specific enough to serve as diagnostic aids.

TECHNOLOGY AND THE DIAGNOSIS OF SCHIZOPHRENIA: THE SEARCH FOR A BIOLOGICAL MARKER

Recent advances in technology, notably in computerized brain imaging, have opened up new possibilities in psychiatry. It is now theoretically possible to visualize ever smaller structural idiosyncracies in the brain, to compare brain structures within and among diagnostic groupings and, with Positron Emission Tomography, to watch the brain at work, to visualize where and with what avidity specific radioactively labelled molecules--glucose for instance, or dopamine--are utilized.

Specifically in relation to schizophrenia, new advances in chromatography, receptorology, virology, immunology, neuropsychology, electroencephalography, and oculography, as well as in imaging, have extended the exploration of theories of etiology and have fueled the continuing search for a biological marker.

Diagnosis and Schizophrenia

Before describing some of these advances, I will summarize the current state of diagnosis in schizophrenia. Clinical diagnosis has become more rigorous over the last five years. Operationalized criteria for inclusion and exclusion have made the diagnosis of schizophrenia more reliable. The chances are better now that two clinicians, interviewing the

same patient, will agree on diagnosis (1,2). Standard structured interviews used for research purposes elicit the necessary criteria more uniformly and further increase the diagnostic reliability (3).

It is the lifetime diagnosis thus arrived at that is conceptualized as a distinct biological entity for which biological markers can be sought. When it comes to schizophrenia, this is an oversimplification. An RDC or DSM-III or any other operationally defined schizophrenia is basically a grouping by behavioural symptoms. These symptoms lack specificity and have little stability in time. The narrower the criteria used, the better the reliability but the worse, seemingly, the heritability (4,5)--i.e., whatever is supposedly inherited does not correlate with symptoms.

The notion that the current diagnostic category of schizophrenia represents a homogeneous biological entity cannot seriously be upheld. Even the broad distinction between schizophrenia and affective psychosis, while it can be made reliably cross-sectionally, becomes blurred over time and in many families the two types of psychosis appear not to "breed true" (6,7). There is a recent report in the literature of identical triplets, two diagnosed schizophrenic and one bipolar affective psychosis by a variety of diagnosticians using a variety of interview schedules and criteria (8). There is another recent report on the psychotic manifestations of Huntington's Disease, a proven biological entity affecting known brain sites. The report tabulates how markedly the psychotic symptoms vary from person to person in response, presumably, to very similar brain pathology (9). In other words, for psychiatric disease, there seems to be no exact correlation between a homogeneous symptom picture and a homogeneous bioetiological process.

To search for a common biological marker in everyone who fulfills certain criteria for schizophrenia is naive. It is perhaps somewhat less naive to look at those with specific symptom clusters (i.e., paranoid vs. non-paranoid or positive vs. negative syndrome)(10). An even better strategy is to look for a common marker in a family where several members are psychotic (11,12). This may even lead, as in Huntington's (13), to the identification of a chromosomal marker.

For the purposes of this discussion, I am going to select four current hypotheses of causation in schizophrenia and evaluate the results of new findings in the light of the four theories.

The Dopamine Hypothesis

The dopamine hypothesis, perhaps the best known, postulates that schizophrenia may be a result of a functional excess of dopamine in parts of the brain (14). A variant of this hypothesis is that schizophrenia is actually a <u>hypo</u>dopaminergic disease, compensated for by supersensitive dopamine receptors, thus producing a functional <u>hyper</u>dopaminergic state which cannot be measured by the usual indices of neurotransmitter excess (14).

A. Dopamine Metabolites

If there were a true dopamine excess, increased levels of dopamine metabolites should be evident peripherally. Gas chromatography/mass spectroprotometry and high performance liquid chromatography techniques allow for the analysis of dopamine dehydrophenylacetic acid (DOPAC) and homovanillic acid (HVA). Consistently higher (or lower for the hypodopaminergic hypothesis) levels in schizophrenic subjects would support the hypothesis. The literature is inconsistent (15). An interesting attempt to look at schizophrenic subjects with a family history, is the recent 25-year follow-up of the Genain quadruplets- four identical sisters, all schizophrenic. Their off-medication DOPAC and HVA levels did not differ significantly from controls although urinary excretion of DOPAC tended to be low for all four quadruplets (16).

B. Dopamine Enzymes

Several enzymes convert dopamine to other molecules. Disproportionately high dopamine activity could result from enzyme deficiencies. Monoamine oxidase (MAO) has been found by several investigators to be significantly lower in platelets from schizophrenic subjects (17) and particularly in those with a family history of illness (18). Several recent studies, however, reviewed by DeLisi et al. (19) have found that low MAO activity is related to neuroleptic treatment. Plasma dopamine-B-hydroxylase (DBH), the enzyme that coverts dopamine to norepinephrine, was found to be consistently lower in all four drug-free Genain quadruplets and their never-medicated family members compared with controls (16). DBH has a strong inherited component and has been found in several studies to be low in schizophrenics but, again, there is the strong possibility that this may be related to medication (19).

C. Dopamine Activity

Eye blink is known to be dopamine mediated and some studies report increased eye blink rates in schizophrenia (20). Dopamine inhibits prolactin and, again, low levels of prolactin have been found to correspond with increased severity in schizophrenia (21). None of these reports has been conclusive enough or dissociated enough from drug effects to be convincing.

Dopamine can be radioactively labelled and visualized by PET scanner. This technique is being developed by Drs. Garnett, Cleghorn, et al. at McMaster (22).

D. Dopamine Receptors

Receptor studies of postmortem brain have shown consistently increased numbers of DA receptors in schizophrenic basal ganglia (23-25). Again, this may be attributable to medication (26). A recent report

suggests that there are two populations of schizophrenics--those in whom receptor numbers are only slightly increased over controls (presumably due to drugs) and those in whom receptor numbers are increased two-fold--some of these subjects known to have never been on drugs (27).

These findings have prompted in vivo studies of DA receptors. There is a neuroleptic binding site on B lymphocytes although it does not appear to have the qualities of a true receptor. Using the same ligand binding assay techniques as in postmortem brain, some groups have reported higher numbers of these high affinity binding sites in lymphocytes of some, particularly paranoid, untreated schizophrenics (28,29).

Positron Emission imaging techniques have been developed to count brain DA receptors in vivo. A radioactively labelled neuroleptic (methylspiroperidol) is injected intravenously. The numbers of occupied DA receptors are then visualized and counted. The technique is still in its infancy but holds great promise. At present only about 50% of DA receptors are occupied and there is a variability among the controls with respect to sex and age (30). Results in schizophrenia are awaited.

The Norepinephrine Hypothesis

The norepinephrine hypothesis postulates that schizophrenia, particularly the paranoid subtype, is related to a chronic excess of brain norepinephrine and perhaps, secondarily, to subsensitive a-adrenergic receptors (31).

Elevations in norepinephrine have been found in the CSF and in postmortem limbic and striatal brain structures of schizophrenic patients (32-34). Altered a-adrenergic receptor function on platelets has been reported and some schizophrenic subjects show subsensitive peripheral a-adrenergic receptors (35). However, the Genain quadruplets and two out of three of their family members had elevated a-adrenergic concentrations in their platelets (16).

In a recent October 1983 debate about the merits of dopamine vs. the norepinephrine hypothesis, Iversen, Reynolds, and Snyder write:

> Changes have been found in levels of dopamine and noradrenaline and/or their metabolites in schizophrenic brain or spinal fluid. However, on balance the changes have been modest and not consistently reproduced in different laboratories. Moreover, in the great majority of instances it has not been possible to rule out the strong likelihood that observed changes are due to treatment with neuroleptics. In the case of the very consistently observed increase in number of dopamine receptors in schizophrenic brain, there is evidence that the changes are related to neuroleptic drug treatment ... Thus, there are no reproducible changes ...

> indicating that either noradrenaline or dopamine is causally involved in the symptoms of schizophrenia ... Dopamine is more likely to be involved than noradrenaline, since L-dopa, which is converted in the brain largely into dopamine with little if any noradrenaline formed, also exacerbates schizophrenic symptoms ... This by no means implies a causal role for dopamine in schizophrenia. However, it does suggest that dopamine systems may be fairly closely linked to whatever is fundamentally aberrant in schizophrenic brain. (36)

Hornykiewicz counters:

> ...there is, on careful examination, less contradiction in reports on brain noradrenaline changes in schizophrenia, or at least its paranoid variety, than there is in reports on dopamine in schizophrenia. At present, therefore, the direct neurochemical evidence for abnormal noradrenaline in schizophrenic brain looks better than any such evidence for a direct brain dopamine disturbance. That is why I find the case of brain noradrenaline more convincing than the case for brain dopamine. (37)

Biochemical theories like the dopamine and norepinephrine hypotheses speak mainly to inherited defects although it is recognized that neurotransmitter activity is influenced both temporarily and more permanently by environmental events. The next two hypotheses, the viral and the specific brain lesion hypotheses, emphasize the environmental, although it is also recognized that inborn vulnerabilities play a role in infection, anoxia, trauma, or faulty learning. The vulnerability marker for schizophrenia, in the light of these more environmental hypotheses, would not be expected to be present in the non-affected relatives of schizophrenic subjects. This, however, has to be tempered by the realization that infections also run in families, that perinatal anoxia and trauma are very much under the influence of social class, and that faulty cognitive strategies may be a family phenomenon.

The Viral Hypothesis

The viral hypothesis postulates an early central nervous system viral or viroid infection to be a necessary antecedent to the development of schizophrenia (38). Herpes, cytomegalic, adeno and rota viruses have been implicated, among others. Antibody titers, done on CSF by enzyme-linked immunosorbent assay (ELISA), have been generally negative but the viral hypothesis has received new impetus by two findings.

The first is the post-mortem report of the architectural disarray of hippocampal cell alignment in schizophrenia (39). The investigators consider the disarray congenital but it could also be secondary to slow virus infection in early life. The hippocampus, incidentally was implicated by Mednick in schizophrenia as a result of the sustained attention and arousal patterns he found in children-at-risk (40). Mednick also pointed out that the hippocampus is especially vulnerable to anoxia, another possible explanation for the histological findings.

The second finding suggestive of a viral cause (although it can be understood in many other ways as well) is the report, confirmed now by many investigators, of enlarged lateral ventricles found in the brains of approximately 20-30% of schizophrenics on CT scan (41).

Those schizophrenics with enlarged ventricles have more defect or negative symptoms, more neuropsychological deficits, poorer premorbid histories and a poorer response to neuroleptic drugs. These are the general findings although they are disputed by some (42,43). There appears to be no association between the size of ventricles and the duration of either illness or treatment although ventricular brain ratio (VBR) increases after middle age presumably because of neuronal atrophy.

In the light of the viral hypothesis, one possible explanation for the increased VBR ratios in a substantial proportion of schizophrenics could be a periventriculitis secondary to cytomegalic virus.

A confounding factor is that VBR's depend on the technique of measurement which is not yet standardized and that ventricular size varies widely in the general population. VBR's tend to run in families, a fact which may still be compatible with the viral hypothesis.

A major problem is specificity in that individuals suffering from many different disorders have now been reported to show increases in VBR (44).

Nuclear Magnetic Resonance, the newer imaging technique, has several advantages over CT for psychiatric research. More planes or section can be visualized, the image depends on the emission of radio frequencies of protons in a magnetic field so there is no X-raying or radioactivity involved and the scan can be safely repeated in longitudinal studies. The contrasts produced by the more powerful NMR's are said to be sharper and clearer than those produced by CT. NMR scanning takes time, however, and requires the subject to sit still over an extended period of time (1-4 hours with breaks). One study using the NMR failed to show qualitative differences between nine schizophrenics and controls in any measure of brain structure (45) but the technique is still new and needs to be quantified.

Recent studies of VBR in twins with schizophrenia shed light on possible genetic/environmental influences on ventricular size. Monozygotic twins in general have very little intra-pair difference in ventricular size, demonstrating a high degree of genetic influence. MZ twins

discordant for schizophrenia show a significantly larger intra-pair difference and those with schizophrenia show consistently larger cerebral ventricles than both their non-schizophrenic co-twins and normal controls (46). This suggests an additional environmental process in the enlarged ventricles found in schizophrenia. In a group of schizophrenics of twin birth, no evidence of ventricular enlargement was found when there was a family history of major psychiatric disorder. Among those schizophrenics without such a family history, cerebral ventricular size was significantly increased and there was also evidence of birth complications. Among normal control twins, those who reported complicated births had significantly larger ventricles (47). These investigators' conclusion is that ventricular enlargement may be a non-specific indicator of a past or present cerebral insult which produces psychotic illness (not just schizophrenia) in those with relatively low genetic predisposition for the illness. Those with high predisposition develop the illness presumably with or without the addition of cerebral pathology. It appears then, that increased VBR is mainly an indicator of environmental brain insult and that it correlates well with obstetric complications (48).

The Specific Brain Lesion Hypothesis

The word "lesion" in the specific brain lesion hypothesis is used loosely to refer to a regional deficit or abnormality as opposed to a total brain dysfunction. In this hypothesis, the actual cause of the "lesion" (whether neurochemical or viral or anoxic) is less important than its site, site presumably dictating brain function or dysfunction. In other words, a specific brain site or sites when disrupted in any number of ways could result in a schizophrenic syndrome. State or trait markers are then looked for in the structure or function of the brain region thought to be dysfunctional. The hippocampus has come under suspicion, as stated above, because of attentional problems common to many schizophrenics. The limbic lobe is most often implicated and the striatum has been suspect as well. Recent technology has advanced two new regional hypotheses: hypofrontality and left hemispheric dysfunction.

A. Hypofrontality

Both blood flow and glucose uptake measures identify the site of brain work. Blood flow studies have reported a relative decrease in cerebral blood flow in the frontal lobes in schizophrenia (49,50). Buchsbaum et al. (51) and Farkas et al. (52) have confirmed a relative hypofrontality in drug-free schizophrenics compared to controls on cerebral glucography or radioactive glucose use as measured by PET. Buchsbaum et al. have a new report with an increased sample and under sensory stimulation conditions (in order to prevent uncontrolled psychological activity). They have studied 16 drug-free schizophrenics, 11 affective disorder patients and 19 controls. They confirm a lower anterior/posterior gradient, i.e., a relative hypofrontality in schizophrenia than in normal controls but there are no differences between schizophrenic and affective disorder subjects. The

investigators suggest several possible explanations for this, among them the possibility that normal activity in frontal areas is responsible for planning goal-directed behaviour which is disrupted in both the major psychoses (53).

Relative hypofrontality in schizophrenia is consistent with reports of slow frontal EEG activity in a topographic study of schizophrenics and controls (54).

It is probably also consistent with eye tracking dysfunction (ETD) in schizophrenia, one of the most widely confirmed abnormalities in schizophrenia research. ETD's consist of larger than expected saccadic intrusions in smooth eye pursuit, the tracking of a smooth moving target such as a swinging pendulum, as measured by electrooculography. ETD's are not affected by severity of illness nor by neuroleptic treatment. They are found in 50-85% of schizophrenics, 8% of normals, and 40% of affective disorder patients (although they do appear to be influenced by lithium). Thirty-four percent of parents of schizophrenics versus 10% of parents of affective disorder patients show ETD's. They are thought to be due to the disinhibition of frontal eye field mechanisms and may be a vulnerability marker for psychosis, if not for schizophrenia (55,56). A recent report suggests they are a marker for schizotypal personality disorder, a possible precursor of schizophrenia (57).

B. Left Hemisphere Impairment

Several lines of investigation point to the specific involvement of the left hemisphere in schizophrenia.

a) Neuropsychological testing and event-related potentials

Neuropsychological tests that can localize dysfunction (certain subtests from the Halstead Reitan Battery and the Luria-Nebraska) implicate a general cognitive disorder in schizophrenia with a relatively greater impairment of left hemisphere tasks, especially left tempero-parietal (58). This finding, however, is inconsistent (59). It is nevertheless supported by evoked potential evidence showing smaller amplitude event-related potentials (ERP) following a sensory stimulus in schizophrenics versus controls. This is most marked at about 300 milliseconds after the stimulus (the P300 peak) over the tempero-parietal area, especially on the left. The timing and location suggest an impairment of cognitive strategies (which could be a result of faulty learning) in response to non-specific stimuli (60).

b) Psychophysiology

Electroencephalography (EEG)(61), blood flow (62), brain density on CT scan (63,64) and Buchsbaum et al.'s original (51) but not more recent (53) PET glucography studies support greater left/right hemisphere impairment in schizophrenia as compared to controls. Three of the Genain quadruplets had lower glucose uptake on the PET scan on the left than on the right, a reversal of the normal pattern (65).

c) Perceptual tasks

Skin conductance response (SCR)(66), dichotic listening tasks (67), visual tachistoscopic tasks (68) and lateralized tactile tasks (69), taken together, also support a left hemisphere disadvantage in schizophrenia.

d) Motor tasks

Handedness (70,71) and lateral eye movement tasks (72) have yielded more conflicting results.

The difficulties with the interpretation of these accumulating data are many. Variables to be considered in laterality test results are handedness, gender (73), and task difficulty. Different sensory modalities tap different brain regions which makes generalization of results across modalities impossible (74). Laterality is state-dependent to a large extent (75) and medication effects have to be taken into account (76). Some results, furthermore, apply only to subgroups of schizophrenics (77).

Nevertheless, it is tempting to try to integrate what is known about antero/posterior differences and left/right differences in schizophrenics versus controls with what is known about neuro-chemical asymmetries. Dopamine activity appears to be asymmetrically distributed in the brain (78,79) and there appear to be antero/posterior gradients in norepinephrine distribution (80). It may be that for specific subgroups of schizophrenics, neurochemical and regional brain etiological hypotheses will prove mutually confirmatory.

Summary

In summary, new technical advances have made it possible to explore current hypotheses about the origins of schizophrenia. Markers of dopamine and norepinephrine activity, brain atrophy, relative hypofrontality, smooth eye pursuit dysfunction and relative left-hemisphere impairment all seem promising but are, as yet, not specific enough to serve as diagnostic aids.

REFERENCES

1. Endicott, J., Nee, J., Fleiss, J., Cohen, J., Williams, J.B.W. and Simon, R. Diagnostic criteria for schizophrenia: Reliabilities and agreement between systems. Arch Gen Psychiat, 1982, 39, 884-889.

2. Williams, J.B.W. and Spitzer, R.L. Research diagnostic criteria and DSM-III, an annotated comparison. Arch Gen Psychiat, 1982, 39, 1238-1289.

3. Hesselbrock, V., Stabenau, J., Hesselbrock, M., Mirkin, P. and Meyer, R. A comparison of two interview schedules. Arch Gen Psychiat, 1982, 39, 674-677.

4. Pope, H.G. Jr., Jonas, J.M. and Cohen, B.M. et al. Failure to find evidence of schizophrenia in first-degree relatives of schizophrenic probands. Am J Psychiat, 1982, 139, 826-828.

5. Abrams, R., Taylor, M.A. The genetics of schizophrenia: A reassessment using modern criteria. Am J Psychiat, 1983, 140, 171-175.

6. Karlsson, J.L. Family transmission of schizophrenia: A review and synthesis. Br J Psychiat, 1982, 140, 600-606.

7. Kendler, K.S. and Hays, P. Schizophrenia subdivided by the family history of affective disorder. A comparison of symptomatology and cause of illness. Arch Gen Psychiat, 1983, 40, 951-955.

8. McGuffin, P., Reveley, A. and Holland, A. Identical triplets: Non-identical psychosis? Br J Psychiat, 1982, 140, 1-6.

9. Caine, E.D. and Shoulson, I. Psychiatric syndromes in Huntington's disease. Am J Psychiat, 1983, 140, 728-733.

10. Dworkin, R.H. and Lenzenweger, M.F. Symptoms and the genetics of schizophrenia: Implications for diagnosis. Am J Psychiat, 1984, 14, 1541-1546.

11. Baron, M., Gruen, R. and Asnis, L. Schizophrenia: A comparative study of patients with and without family history. Br J Psychiat, 1982, 140, 516-517.

12. Kendler, K. and Hays, P. Familial and sporadic schizophrenia: A symptomatic, prognostic and EEG comparison. Am J Psychiat, 1982, 139, 1557-1562.

13. Gusella, J.F., Wexler, N.S. and Conneally, P.M. et al. A polymorphic DNA marker genetically linked to Huntington's disease. Nature, 1983, 306, 234-238.

14. Haracz, J.L. The dopamine hypothesis: An overview of studies with schizophrenic patients. Schiz Bull, 1982, 8, 438-469.

15. Gattaz, W.F., Waldmeier, P. and Beckmann, H. CSF monoamine metabolites in schizophrenic patients. Acta Psychiat Scan, 1982, 66, 350-360.

16. DeLisi, L.E., Mirsky, A.F. and Buchsbaum, M.S. et al. The Genain quadruplets 25 years later: A diagnostic and biochemical follow-up. Psychiat Research, 1984, 13, 56-76.

17 DeLisi, L.E., Phelps, B.H. and Bridge, T.P. et al. MAO and schizophrenia. In Hanin, I., Usdin, E. (eds.) Biological Markers in Psychiatry and Neurology. New York: Pergamon Press, 1982.

18. Baron, M. and Levitt, M. Platelet monoamine oxidase activity: Relation to genetic factors in schizophrenia. Psychiat Res, 1980, 3, 69.

19. DeLisi, L.E., Wise, C.D. and Bridge, T.P. et al. A probable neuroleptic effect on platelet monoamine oxidase in chronic schizophrenic patients. Psychiat Res, 1981, 4, 95.

20. Stevens, J.R. Blink rates: Psychosis or tardive dyskinesia? Am J Psychiat, 1978, 135, 223-227.

21. Kleinman, J.E., Weinberger, D.R. and Rogol, A.D., et al. Plasma prolactin concentrations and psychopathology in chronic schizophrenia. Arch Gen Psychiat, 1982, 39, 655-657.

22. Garnett, E.S., Firnau, G. and Nahmias, C. Dopamine visualized in the basal ganglia of living man. Nature, 1983, 305, 137-138.

23. Lee, T., Seeman, P. and Tourtellotte, W.W., et al. Binding of ^{3}H-apomorphine in schizophrenic brains. Nature, 1978, 274, 897-900.

24. Lee, T. and Seeman, P. Elevation of neuroleptic/dopamine receptors in schizophrenia. Am J Psychiat, 1980, 137, 191-197.

25. Owen, F., Cross, A.J. and Crow, T.J., et al. Increase dopamine receptor sensitivity in schizophrenia. Lancet, 1978, 2, 223-226.

26. MacKay, A.V.P., Bird, E.D. and Spokes, E.G., et al. Dopamine receptors and schizophrenia: Drug effect or illness? Lancet, 1980, 2, 915-916.

27. Seeman, P., Ulpian, C. and Bergeron, C., et al. Bimodal distribution of dopamine receptor densities in brains of schizophrenics. Science, 1984, 225, 728-731.

28. LeFur, G., Zarifian, E. and Phal, T., et al. (^{3}H)-spiroperidol binding on lymphocytes: Changes in two different groups of schizophrenic patients and effect of neuroleptic treatment. Life Sciences, 1983, 32, 249-255.

29. Rotstein, E., Mishra, R.K. and Singal, D.P., et al. Lymphocyte ^{3}H-spiroperidol binding in schizophrenia. Preliminary Findings. Prog Neuro Psychopharm & Biol Psychiat, 1983, 7, 729-732.

30. Wong, D.F., Wagner, H.N. and Dannals, R.F., et al. Effects of age on Dopamine and Serotonin receptors measured by Positron Tomography in the living human brain. Science, 1984, 226, 1393- 1396.

31. VanKammen, D.P. Mini-review. Impaired noradrenergic transmission in schizophrenia. Life Sciences, 1984, 34, 1403.

32. Lake, C.R., Steinberg, D.E. and van Kammen, D.P., et al. Schizophrenia: Elevated cerebrospinal fluid norepinephrine. Science, 1980, 207, 331-333.

33. Farley, I.J., Price, K.S. and McCullough, E., et al. Norepinephrine in chronic paranoid schizophrenia: Above normal levels in limbic forebrain. Science, 1978, 200, 456-458.

34. Kemali, D. and Del Vecchio, Maj. M. Increased noradrenaline levels in CSF and plasma of schizophrenic patients. Biol Psychiat, 1982, 17, 711-718.

35. Kafka, M.S., van Kammen, D.P. and Kleinman, J.E., et al. Alpha-adrenergic receptor function in schizophrenia, affective disorder, and some neurological diseases. Communications in Psychopharmacology, 1980, 4, 477.

36. Iversen, L.L., Reynold, G.P. and Snyder, S.H. Pathophysiology of schizophrenia--causal role for dopamine or noradrenaline. Nature, 1983, 305, 577.

37. Hornykiewicz, O. Reply to Iverson et al. Nature, 1983, 305, 577-578.

38. Torrey, E.F. and Peterson, M.R. The viral hypothesis of schizophrenia. Schiz Bull, 1976, 2, 136.

39. Scheibel, A.B. and Kovelman, J.A. Disorientation of the hippocampal pyramidal cell and its processes in the schizophrenic patient. Biol Psychiat, 1980, 16, 101-102.

40. Mednick, S.A. Breakdown in individuals at high risk for schizophrenia: Possible predispositional perinatal factors. Mental Hygiene, 1970, 54, 50-63.

41. Maser, J.D. and Keith, S.J. CT scans and schizophrenia--report on a workshop. Schiz Bull, 1983, 9, 265-283.

42. Jernigan, T.L., Zatz, L.M. and Moses, J.A., et al. Computed tomography in schizophrenics and normal volunteers. Arch Gen Psychiat, 1982, 39, 765-770.

43. Nasrallah, H.A., Jacoby, C.G. and McCalley-Whitters, M., et al. Cerebral ventricular enlargement in subtypes of chronic schizophrenia. Arch Gen Psychiat, 1982, 39, 774-777.

44. Rieder, R.O., Mann, L.S. and Weinberger, D.R., et al. Computed tomographic scans in patients with schizophrenia, schizoaffective and bipolar affective disorder. Arch Gen Psychiat, 1983, 40, 735-739.

45. Smith, R.C., Calderon, M. and Ravichandran, G.K., et al. Nuclear magnetic resonance in schizophrenia: A preliminary study. Psychiat Res, 1984, 12, 137-147.

46. Reveley, A.M., Reveley, M.A. and Clifford, C.A., et al. Cerebral ventricular size in twins discordant for schizophrenia. Lancet i, 1982, 540-541.

47. Reveley, A.M., Reveley, M.A. and Murray, R.A. Cerebral ventricular enlargement in non-genetic schizophrenia: A controlled twin study. Br J Psychiat, 1984, 144, 89-93.

48. Schulsinger, F., Parnas, J. and Petersen, E.T. et al. Cerebral ventricular size in the offspring of schizophrenic mothers. Arch Gen Psychiat, 1984, 41, 602-606.

49. Ingvar, D.H. and Franzen, G. Abnormalities of cerebral blood flow distribution in patients with chronic schizophrenia. Acta Psychiat Scan, 1974, 50, 425-462.

50. Ariel, R.N., Golden, C.J. and Berg, R.A. et al. Regional cerebral blood flow in schizophrenics. Arch Gen Psychiat, 1983, 40, 258-263.

51. Buchsbaum, M.S., Ingvar, D.H. and Kessler, R. et al. Cerebral glucography with positron tomography. Arch Gen Psychiat, 1982, 39, 251-259.

52. Farkas, T., Wolf, A.P. and Jaeger, J. et al. Regional brain glucose metabolism in chronic schizophrenia. Arch Gen Psychiat, 1984, 41, 293-300.

53. Buchsbaum, M.S., DeLisi, L.E. and Holcomb, H.H. et al. Anteroposterior gradients in cerebral glucose use in schizophrenia and affective disorders. Arch Gen Psychiat, 1984, 41, 1159-1166.

54. Morihisa, J.M., Duffy, F.H. and Wyatt, R.J. Brain electrical activity mapping (BEAM) in schizophrenic patients. Arch Gen Psychiat, 1983, 40, 719-728.

55. Holzman, P.S., Solomon, C.M. and Levin, S. et al. Pursuit eye movement dysfunctions in schizophrenia. Arch Gen Psychiat, 1984, 4, 136-139.

56. Levin, S. Frontal lobe dysfunctions in schizophrenia I. Eye movement impairments. J Psychiat Res, 1984, 18, 27-55.

57. Siever, L.J., Coursey, R.D. and Alterman, J.S. et al. Impaired smooth pursuit eye movement: Vulnerability marker for schizotypal personality disorder in a normal volunteer population. Am J Psychiat, 1984, 14, 1560-1566.

58. Taylor, M.A., Greenspan, B. and Abrams, R. Lateralized neuropsychological dysfunction in affective disorder and schizophrenia. Am J Psychiat, 1979, 136, 1031-1034.

59. Kolb, B. and Whishaw, I.Q. Performance of schizophrenic patients on tests sensitive to left or right frontal, temporal or parietal function in neurological patients, J Nerv Ment Dis, 1983, 171, 435-443.

60. Roth, W.T., Pfefferbaum, A. and Kelly, A.F. et al. Auditory event-related potentials in schizophrenia and depression. Psychiat Res, 1981, 4, 199-212.

61. Abrams, R. and Taylor, M.A. Differential EEG patterns in affective disorder and schizophrenia. Arch Gen Psychiat, 1979, 36, 1355-1358.

62. Franzen, G. and Ingvar, D.H. Abnormal distribution of cerebral activity in chronic schizophrenia. J Psychiat Res, 1975, 12, 199-214.

63. Golden, C.J., Graber, B., Coffman, J., Berj, R., Block, S. and Brogan, D. Structural brain deficits in schizophrenia. Arch Gen Psychiat, 1981, 38, 1014-1017.

64. Largen, J.W., Calderon, M. and Smith, R.C. Asymmetries in the densities of white and grey matter in the brains of schizophrenics. Am J Psychiat, 1983, 140, 1060-1062.

65. Buchsbaum, M.S., Mirsky, A.F. and De Lisi, L.E. et al. The Genain quadruplets: Electro-physiological, positron emission and X-ray tomographic studies. Psychiat Res, 1984, 13, 95-108.

66. Venables, P.H. The electrodermal psychophysiology of schizophrenics and children at risk for schizophrenia. Controversies and developments. Schiz Bull, 1977, 3, 28-49.

67. Bruder, G.E. Cerebral laterality and psychopathology: A review of dichotic listening studies. Schiz Bull, 1983, 9, 134-151.

68. Gur R.E.: Left hemispheric dysfunction and left hemisphere overactivation in schizophrenia. J Abn Psychol 1978; 87: 226-238.

69. Scarone S, Pieri E. Cazzulo CL: The asymmetric lateralization of tactile extinction in schizophrenia: The possible role of limbic and frontal regions. Br J Psychol 1982; 141: 350-353.

70. Chaugule V.B., Masters R.S.: Impaired cerebral dominance and schizophrenia. Br J Psychiat 1981; 139: 23-34.

71. Piran N, Bigler E.D., Cohen D: Motoric laterality and eye dominance suggest unique pattern of cerebral organization in schizophrenia. Arch Gen Psychiat 1982; 39: 1006-1010.

72. Erlichman H, Weinberger A: Lateral eye movements and hemispheric asymmetry: A critical review. Psychol Bull 1978; 85: 1080-1101.

73. Schweitzer L: Cerebral lateralization: Relation to subject's sex. Cortex 1980; 16: 559-566.

74. Schreider S.J.; Multiple measures of hemispheric dysfunction in schizophrenia and depression. Psychol Med 1983; 13: 287-297.

75. Nasrallah H.A.: Laterality and hemispheric dysfunction in schizophrenia. Dans Henn F.A., Nasrallah H.A. (eds.) Schizophrenia as a Brain Disease, New York: Oxford University Press, 1982.

76. Magaro P.A., Page J: Brain disconnection, schizophrenia and paranoia. J Nerv Ment Dis 1983; 11: 219-227.

77. Gruzelier J.H.: Cerebral laterality and psychopathology: Fact or fiction. Psychol Med 1981; 11: 219-227.

78. Glick S.D., Ross D.A., Hough L.B.: Lateral asymmetry of neurotransmitters in human brain. Brain Res 1982; 234: 53-63.

79. Reynolds G.P.: Increased concentrations and lateral asymmetry of amygdala dopamine in schizophrenia. Nature 1983; 305: 527-529.

80. Morrison J.H., Molliver M.E., Grazanna R: Noradrenergic innervation of cerebral cortex. Widespread effects of local cortical lesions. Science 1979; 205: 313-316.

THE AFFECTIVE DISORDERS

BIOTECHNOLOGICAL ADVANCES IN THE TREATMENT OF AFFECTIVE DISORDERS

BARRY D. JONES, M.D., F.R.C.P.(C)
University of British Columbia
Vancouver

ABSTRACT

Biotechnological advances in the treatment of affective disorders are highlighted. The areas of research examined are:

1) New Treatments
2) Predictors of Treatment Response
3) Monitoring of Treatment, and
4) Theories of Mechanism of Action

The clinical application of research findings is suggested to be disappointing. Problems addressed as obstructive towards further progress in this area include the lack of specificity of action of affective disorder drug treatments and diagnostic heterogeneity.

Biotechnological Advances in the Treatment of Affective Disorders

Cade (1949) in Australia reported the first response of manic patients to lithium. Zeller et al. (1952) found iproniazid to be an effective inhibitor of monoamine oxidase and in 1957 it was found to treat depression. In 1958 imipramine was found to be an effective antidepressant (Kuhn, 1958) and shortly thereafter other tricyclic drugs were introduced. Also, during the 1950's the neuroleptic drugs were being shown to be effective treatments in the control of mania as well as schizophrenic psychoses (Delay and Deniker, 1952; Davis, 1965; Spring et al., 1970). Thus within a 10-year span many of the currently efficacious treatments in the management of affective disorders both unipolar and bipolar were available and the development of these agents heralded not only an era of great clinical advancement in treatment but the potential for further biotechnological advances.

As we enter 1985 approximately 30 years after the development of these pioneering drugs it is interesting to examine the current "state of the art" with respect to biotechnological advances in the treatment of affective disorders. Specifically we need to ask the question, "How far have we advanced, which areas show promise for further advancement, and if advancement is disappointing in some areas, why that might be so?"

Oversimplifying, research in affective disorder treatment can be divided into four sub-headings:

1. The assessment of new treatments;
2. The predictors of treatment response including clinical predictors, laboratory predictors, and pharmacological challenge predictors;
3. Monitoring of treatment;
4. The understanding of the mechanism of action of treatments.

The affective disorder treatments researched can also be divided, again oversimplifying, into three sub-headings:

1. Acute treatment of depression;
2. Acute treatment of mania;
3. Maintenance or illness treatments.

Such a classification allows for a matrix to be formed within which one can examine these particular areas of research and the particular affective disorder treatment that is being studied. This matrix is represented in Figure 1.

As is evident, current research in the area of affective disorders dealing only with those disorders as categorized by DSM III as the major affective disorders and focusing primarily on pharmacological treatments is extensive and deals with a broad range of disciplines and technologies. A complete review of the advances and current research in the treatment of affective disorders is thus beyond the scope of this paper. Instead the purpose of this review will be to highlight those areas of research that are holding promise for progress in the future and secondly to examine the effect of recent and current research on the practical clinical application of treatment of the affective disorders. Finally some discussion of the problems we face in bridging the gap between research and clinical practice will be explored.

Current Research Highlights

New Treatments

Thirty years after the development of the first effective pharmacological treatments in affective disorders, the majority of medications prescribed in 1985 remain the treatments originally developed. Although at first glance this may suggest little progress in the development of new treatments, there have been in fact substantial advances in the clinical armamentarium available to the physician treating affective disorders. With respect to depression, so-called second generation antidepressants (Feighner, 1981) continue to be developed which unlike their tricyclic ancestors are progressively more specific in their action in the brain and as a result demonstrate a side effect profile that is decidedly cleaner relative to the earlier drugs. Also the development of drugs with actions that are atypical, i.e., do not lead to significant reuptake blockade of

noradrenaline or seratonin has led to re-thinking of the putative mechanisms of action of these drugs and in turn the possible etiologies of the illnesses that they treat (Charney et al., 1981). Another advancement has been continued elucidation of potentiation strategies in the treatment of depression. Agents such as thyroid derivatives (Schwarcz et al., 1984; Goodwin et al., 1982) estrogen-progesterone compounds (Oppenheim, 1982; Prange, 1972), transmitter precursors such as tryptophan (Chouinard et al., 1983), tyrosine (Van Praag, 1984), and L-dopa (Baldessarini, 1984) and even the drugs themselves such as lithium (De Montigny et al., 1984; Heninger et al., 1984; Louie, 1984) have been studied and found to offer promise in treating patients whose response to antidepressant drug treatment alone has not been optimal.

In the area of mania what could be termed second generation antimanics have been developed and in the early stages of study promise as do their second generation antidepressant counterparts, the potential for more specific action with less side effects especially compared to neuroleptic drugs. Also these second generation antimanic drugs such as carbamazepine (Ballenger and Post, 1984; Okuma et al.; Post et al., 1984), clonazepam (Chouinard et al., 1983), clonidine (Zubenko et al., 1984), verapamil (Giannini et al., 1984) and valproic acid (Emrich et al., 1984) raise interesting questions concerning mechanism of action of antimanic agents and in turn putative etiologies of this syndrome. Furthermore potentiation strategies continue to be uncovered in the form of thyroid derivatives (Stancer, Persad, 1982; von Hofmann and Bolen, 1969) estrogen progesterone compounds (Oppenheim, 1983), and tryptophan (Chouinard - personal communication).

Maintenance treatment for the affective disorders both unipolar and bipolar have also seen advances most dramatically in the recent appearance of carbamazepine as a proven effective treatment in bipolar disorder and an alternative to lithium therapy (Post et al., 1984). Also potentiation strategies as outlined in episodic treatment of depression and mania may also hold promise as maintenance treatments although these strategies remain still to be tested under double blind controlled conditions.

To summarize, research into the development of new treatments in affective disorders has made progress. However, this progress remains generally in the area of treatments which have fewer side effects or offer atypical modes of action and thus potentially expand the numbers of patients who can respond. Potentiation strategies also hold great promise although research remains at an early stage and reasons for effectiveness of potentiation remain virtually unstudied.

Predictors of Treatment Response

CLINICAL

Research into clinical predictors of treatment response which could potentially assist in first choice of treatment has been carried out for a number of years with unfortunately little success. In the area of depression, divisions such as agitated versus retarded (Bielski and Friedel, 1976; Paykel et al., 1979), psychotic versus non-psychotic (Glassman and Roose, 1981), and typical versus atypical (Quitkin et al., 1979), have not met with consistent predictive value in determining the type of antidepressant therapy or predicting its effectiveness. Similarly clinical variables such as age (Raskin, 1975), rate of onset (Bielski and Friedel, 1976), and concomitant anxiety symptoms such as phobias and depersonalisation (Raskin and Crook, 1976) have not clearly stood the test of controlled scientific study as predictors of treatment response. One relatively consistent finding from a number of studies is the relationship of suicidal behaviour to measures of decreased serotoninergic activity (Asberg et al., 1980; Traskman-Bendz et al., 1980). Possibly with the development of more specific serotoninergic agonist drugs, this clinical variable will predict drug response. Until then, the only clear diagnostic element which alters the treatment of an episode of depression is the differential diagnosis between unipolar versus bipolar illness. Depressed bipolar patients respond with equal efficacy to lithium or standard antidepressant therapy but there is less potential for induction of mania with lithium treatment (Kukopulos et al., 1980; Quitkin et al., 1982).

With respect to manic treatment, clinical predictors of treatment response based on diagnosis and clinical variables remain at early stages of study given that second generation antimanic drugs have only recently been discovered. Post has suggested, based on clinical evidence that carbamazepine may be more effective in rapid cycling manic depressive patients, and in patients with psychotic symptoms, thus mania might be better controlled with this drug in such a patient (Post et al., 1984). However, this will await further research. What does seem to be clear is that manic patients do not need to present with clinical features related to epilepsy within their manic episodes to be responders to carbamazepine nor do such patients in fact respond better to carbamazepine (Post et al., 1984). Clinical predictors for other antimanic drugs including lithium have not been or remain to be elucidated.

Concerning illness or maintenance treatment of affective disorders, diagnosis and clinical variables as predictors of treatment response parallel to some degree what has been described previously regarding the differentiation between unipolar and bipolar depression. Bipolar illness certainly requires maintenance treatment with either lithium or carbamazepine. Unipolar illness is usually treated with maintenance antidepressant treatment although studies suggest that lithium can be as effective (Angst et al., 1978; Schou et al., 1979; Prien et al., 1980). As stated previously concerning mania, carbamazepine may be more clinically effective for rapid cycling illness of a bipolar type given that lithium is

notoriously ineffective. However potentiation strategies using lithium as the main treatment remain an interesting alternative to carbamazepine in the treatment of such conditions.

LABORATORY

Since the development of the antidepressant and antimanic drugs thirty years ago there have been numerous attempts to find laboratory measurements which predict treatment response or drug of first choice. In the area of depression, the major thrust of investigation for a number of years has been the measurement of neurotransmitter metabolites such as 5HIAA-metabolite of serotonin (Van Praag, 1977; Asberg et al., 1973), MHPG-metabolite of norepinephrine (Maas et al., 1972; Schildkraut, 1973; Beckmann and Goodwin, 1975; Hollister et al., 1980) and HVA-metabolite of dopamine (Jimerson and Post, 1984) and to attempt to correlate deficiencies in these matabolites with drug efficacy based on the drug's potential for increasing activity of the transmitter putatively deficient. The strategy has led to a number of findings which although reaching significance statistically in some studies have met with non-significant findings in other studies (Spiker, 1984). However non-significant findings might be expected given the noise in the system, i.e., the inability of any given metabolite to truly reflect central activity of the neurotransmitter, the possibility that the drug specificity measured is not involved with the action of the antidepressant or that the drug is not in fact specific for that given activity. One finding that has stood the test of time better than others in this area has been that low pretreatment levels of urinary MHPG correlate with favourable response to relatively specific noradrenergic reuptake blocking drugs such as desipramine (Maas et al., 1972; Schildkraut, 1973; Beckmann and Goodwin, 1975; Hollister et al., 1984). However, findings that urinary MHPG reflects as little as 20% of central norepinephrine turnover (Blombery et al., 1980) leave the data intact but the interpretation uncertain. Measurements such as red blood cell COMT (Davidson et al., 1976), an important metabolic enzyme for catecholamines, and psychophysiological measurements such as average evoked response to visual stimulation (Buchsbaum, 1984) have been suggested to identify imipramine responders. More recently techniques such as receptor binding using labelled antidepressant ligands (Hrdina, 1983) have been thought to hold promise in identifying specific sub-types of depression (imipramine binding = serotoninergic reuptake sites; desipramine binding = noradrenergic reuptake sites) and thus possibly drug response. These techniques still remain in their earlier stages of development.

In the treatment of mania, laboratory measurements predicting response extend beyond those used in depression into areas such as membrane characteristics. In this regard lithium has been hypothesied to be transported through ion carrying channels of cell membranes (Dorus et al., 1983). Certain characteristics of this transportation mechanism have been thought to predict lithium response (Pandey et al., 1979).

Concerning maintenance or illness treatments, laboratory predictors of treatment response are generally considered to be those that might predict response to antidepressant or antimanic therapies. In general, however, research in this area has been limited.

PHARMACOLOGICAL CHALLENGE

Difficulties finding clear relationships between drug response and clinical or laboratory findings has led more recently to the strategy of pharmacological challenge. The model for this approach is that differences not evident in baseline state may emerge when patients are given a single dose of a drug or agent that affects a particular neurotransmitter system. The response to such a challenge could then be used to predict antidepressant, antimanic or maintenance treatment response.

In the treatment of depression, the prototype of such challenge studies was that carried out using amphetamine response of mood improvement as a predictor of patients who would respond to imipramine or desipramine (Fawcett and Siomopoulos, 1971).

Amphetamine has also been used as a neuroendocrine challenge using growth hormone response as a predictor of noradrenergic receptor sensitivity (Sanger et al., 1976). Within the same paradigm clonidine has been used to identify putative subsensitive noradrenergic mechanisms in depressed patients (Matuesek et al., 1980). While not technically a drug challenge, the dexamethasone suppression test (Brown, 1984) and the TRH-TSH test (Extein et al., 1984) have received a lot of attention in recent literature as not only diagnostic aids but also predictors of treatment response. It has been contended that dexamethasone escape reflects a deficit of noradrenergic function with those showing escape responding to noradrenergic drugs such as desipramine (Brown et al., 1980). However, the specificity of the dexamethasone suppression test has more recently been questioned not only with regard to drug response but also as a diagnostic aid (Hwu et al., 1981) and similar difficulties exist with the TRH-TSH test and the other tests previously mentioned (Insel and Murphy, 1984). Finally antidepressants themselves in terms of early treatment changes such as REM suppression have been suggested as a predictor of therapeutic response (Kupfer et al., 1976).

In the treatment of mania, pharmacological challenges to predict drug response have followed similar patterns to those attempts to predict drug response within the area of depression. In addition the use of early treatment changes such as decrease of thyroxine levels after carbamezapine therapy as a predictor of carbamezapine response have been suggested (Post et al., 1984). However these findings, as with the finding of early REM suppression predicting antidepressant response may simply reflect the degree to which a drug is available to act, whether that be to produce side effects or therapeutic effect. Still if such findings are predictive of long-term outcome on treatment they are useful given the delayed response to treatment with these agents.

Such findings are especially relevant to the issue of maintenance or illness treatment and predictors of treatment response. Such tests as DST and its normalization in previous patients demonstrating lack of suppression has been suggested to predict not only treatment of choice but also the point of time when treatment within maintenance therapy can be discontinued (Brown, 1984). However previous considerations concerning the specificity and usefulness of the DST and other tests must be kept in mind in this area as well.

Treatment Monitoring

Research in the area of treatment monitoring has for the most part taken the form of plasma level measurements of drugs in correlation with clinical response. In the treatment of depression the earliest literature in this regard suggested that for antidepressants such as nortryptiline there was a therapeutic window within which treatment was effective but beyond which efficacy was lost (Potter and Sinnoila, 1984). Such a therapeutic window effect has not been found for all antidepressants and remains most well substantiated for nortryptiline. However numerous methodological problems in studies attempting to find correlations of plasma level measurement with clinical response may have obscured possible relationships. New techniques in monitoring treatment without directly measuring plasma levels may hold promise in the future. Measurements of drug action such as displacement from receptors of labelled ligands or transmitter uptake inhibition in peripheral membranes such as platelets have been suggested to correlate with clinical response (Singjaerde, 1984). These techniques however will await further research before replacing other techniques of treatment monitoring in depression.

With respect to mania and treatment monitoring the use of lithium levels to monitor acute management of mania is well established. Evidence in this area suggests that levels of lithium up to 1.5 milliequivalents per litre but not above are effective in the control of acute mania (Baldessarini, 1979). However, ultimately these levels are guidelines and do not replace the clinical monitoring of side effects in the patient and adjustment of dose accordingly.

With respect to illness treatment similarly plasma levels of lithium remain well established in the clinical armamentarium of the treatment and maintenance of manic depressive illness. However, extensive research corroboration of the levels agreed upon by most physicians as therapeutic does not exist and in fact some studies would suggest lower levels as being more efficacious than the ones accepted by clinical lore in both the treatment of acute mania and maintenance treatment of bipolar illness (Hullin, 1980). Plasma levels of carbamazepine are in the process of being established with respect to therapeutic efficacy. At the present time most clinicians are using levels which are considered to be similar to therapeutic levels in the control of various forms of epilepsy. One final area of research concerns the measurement of drugs such as lithium in red blood cells or in saliva. Intracellular red blood cell measurements and resultant intra-extracellular ratios of lithium may reflect a sub-group of

patients with high ratios who have manic-depressive illness and another group who are particularly sensitive to lithium toxicity (Cooper and Carroll, 1981). Saliva lithium measurements would be less invasive than plasma level measurements but difficulties arise with the finding that ratio of plasma level to saliva level varies greatly between individuals (Sims, 1980). In general the clinical usefulness of both these techniques remains questionable.

Mechanism of Action

The study of the mechanism of action of agents within the affective disorders is an active research area and it is hoped with such understanding will come similar understanding of the etiology and pathophysiology of the disorders themselves. In the area of depression much new research has evolved with the development of receptor binding techniques and this has led to a number of new hypotheses concerning the mechanism of antidepressant action. Previous suppositions of antidepressant efficacy through reuptake blockade of relevant neurotransmitters such as serotonin and noradrenalin have given way to receptor change theories. These theories are more attractive in that the time course of response to antidepressants being delayed corresponds to the gradual changes in receptor activity whereas reuptake blockade occurs immediately after antidepressants are started. The literature concerning receptor change is a complex one with as many receptors incriminated as drugs are available to treat the disorder. However patterns of findings would seem to suggest that many antidepressants exert their effect through the development of subsensitive beta-adrenergic postsynaptic receptors and/or supersensitive serotonergic postsynaptic receptors (Charney, 1981). This in turn opens the door to speculation that depression is the result of overactivity of the beta-adrenergic system whereas in other cases it may be due to decreased activity of serotonergic systems or finally a combination of both. A unitary hypothesis of how all antidepressants work although striven for by some theoreticians is not necessary if one accepts the distinct possibility that depression is due to multiple etiologies.

Research into the area of mechanism of action of drugs efficacious for mania has taken a slight turn recently with the development of the second generation antimanic agents. Since a number of these drugs are also antiepileptic, drug theories concerning a link between mania and epileptic phenomena have been put forward. One suggestion is that some drugs effective in the treatment of mania such as carbamazepine and valproic acid, have their effect through their ability to suppress kindling within the temporal lobe area (Post et al., 1984). Interestingly electroconvulsive therapy has also been found to suppress kindling and seizure potential although in itself it is a seizure producing treatment. The efficacy of clonazepam in mania further supports this hypothesis (Chouinard et al., 1983). However a lack of correlation between efficacy and the ability to suppress kindling when including drugs such as phenytoin raises doubts about this hypothesis. Similarly the oldest antimanic drug lithium, and other drugs with efficacy in mania such as neuroleptics do not seem to have their mechanism

of action in mania explainable in the same way. Thus other theories such as membrane effects, receptor changes and neurotransmitter alterations continue to receive attention (Bunney and Garland, 1984).

Theories concerning mechanism of action of treatments in the maintenance of affective disorders basically fall within the same categories as those that look at efficacy in episode treatments except that an effective maintenance treatment must contain within its explanation for mechanism of action the ability to suppress both depression and mania. This poses some difficulties in theory construction. Although one explanation would be the ability of prophylactic drugs to suppress receptor supersensitivity (mania) in response to decreased activity of a neurotransmitter system (depression), thus preventing a switch and controlling oscillation of neurotransmitter activity, such theories remain very speculative at this time (Bunney and Garland, 1984).

Current Clinical Applications

The current clinical applications of much of the research that has occurred over the last thirty years and most recently in the area of treatment of affective disorders is presented in Figure 2. As can be seen, little from Figure 1 in terms of current research interests and research which has gone before it has led to actual practical changes in the management and treatment of patients with affective disorders. The major area of advancement is within the category of new treatments. As already stated the development of second generation antidepressant and antimanic drugs and the arrival of carbamazepine and various potentiation strategies for both episodic and illness treatments have opened up the options for the clinician in treating these disorders. Furthermore the new agents in many cases possess not only different modes of action but also a cleaner profile of side effects which allows for more patient comfort and possibly more importantly better compliance with treatment. However it should be stated that despite these advances drastically new highly efficacious treatments have not supplanted the original medications introduced some thirty years ago.

Further disappointment occurs when examining other categories of research and their effect on practical management of patients. As can be seen in Figure 2 only plasma level measurement of lithium in the treatment of mania and maintenance treatment of bipolar, occasionally unipolar illness, is actively pursued as a day-to-day part of clinical practice. Also only diagnosis between unipolar and bipolar affective disorder is used as a predictor of treatment response and drug of choice in maintenance treatment. An obvious question that arises is the reason for thirty years of research resulting in so little change in practical management of patients with affective disorders especially in areas such as predictors of treatment response and treatment monitoring. Although there are a multitude of possible explanations, certain reasons that are possibly within our control to change can be examined. One of the most obvious difficulties in translating research activity into practical improvements in patient management

and treatment is poor methodology. Certainly much research that is conducted suffers from methodological weaknesses that make the research virtually uninterpretable and/or inapplicable. Linked to this difficulty is the problem of inadequate instrumentation. Instrumentation used in the study of affective disorder treatments must overcome the inaccessibility of the brain to invasive techniques and present as little "noise" as possible to minimize what is already a study system full of variables that cannot be controlled. However with time as research methodologies are improved and investigators have available instrumentation which is "state of the art," we will be less the victim of these difficulties in applying our knowledge to practice. As a result, other potential problems such as diagnostic heterogeneity can be broached. An implicit assumption made in much of the research done in the area of affective disorder treatment is that these disorders have few etiologies. However if we draw analogies to other areas of medicine that deal with syndromes affecting large percentages of the population it is most probable that the major affective disorders are not simply two disorders, unipolar and bipolar illness but multiple illnesses with multiple etiologies resulting in the end point production of depression and depression alternating with mania. This may be seen as a reason for difficulties in establishing productive research given that at any given point in time mixed etiologies are being studied even using the most rigidly defined operational criteria for diagnosis.

Strategies such as laboratory measurements or pharmacological challenges attempting to predict treatment response can only fail given the pathophysiological mechanisms they are attempting to uncover might be present in only a small number of patients studied. Treatment monitoring could also be affected given that different drug levels or drug activity could be necessary depending on the specific etiology for the given illness treated. Linked with the difficulty of diagnostic heterogeneity in allowing for progress in research in the treatment of affective disorders is the lack of availability of drugs which are specific in their action as effective treatments for these illnesses. Drawing an analogy to anaemia, the most effective symptomatic treatment of anaemia whatever the cause, is replacement of blood. However, the use of an agent such as B12 injections would, in a group of 100 patients presenting to a clinic with anaemia, possibly improve only one. Without predictors of treatment response such as the clinical findings of neurological deficits and laboratory results such as macrocytic anaemia and decreased B12 levels, such an agent would be quite possibly missed as an effective treatment. This would in turn lead to the inability to discover that some anaemias are related to B12 deficiency, that the mechanism of action of B12 is to replace that deficiency, and that there are clinical and laboratory correlations that would identify such patients affected. In the area of affective disorder treatment, up until recently we have been in a position of treating affective disorders with the equivalent of blood transfusion, in terms of the multiple actions of antidepressants such as trycyclics, MAO inhibitors, electroconvulsive therapy and lithium. Indeed by the nature of the criteria that are used by governments and drug companies that lead to a drug being studied and eventually marketed, there is a bias towards continuing to find non-specific broadspectrum approaches to the treatment of affective disorders. If we add to this the tendency to study short-term state episodic treatments

compared to long-term trait illness treatments, we can again see that the tendency analogous to the effectiveness of a blood transfusion in the treatment of anaemia would be to find non-specific treatments that work but may not maintain a patient's remission over an extended period of time because the primary pathophysiological mechanism has not been touched.

How are we then to apply ourselves to these difficulties in order to improve the applicability and productiveness of research in this area? First and foremost it is necessary to search for more specific drugs not just in terms of their ability to produce less side effects but more hopefully in their ability to be specific in their actions in the brain. However, specific for what? The answer to this question is speculative with respect to relevant central pharmacological effects. However clinically relevant specificity can be measured. Analogous again to the discussion of anaemia and B12 the most interesting drug although perhaps not the most commercial would be one which worked in otherwise drug unresponsive patients but which had little efficacy in the majority of patients and a drug which more clearly than other drugs showed an all-or-none response. Furthermore the drug ideally would treat both depression and mania, would not induce mania in unipolar patients, and would be a drug that would work long term. To study such drugs would obviously, based on these criteria, be expensive. However, if such agents could be found which show this profile consistently in a small percentage of patients, then these patients could be further studied for clinical, laboratory or pharmacological correlations in order to hopefully elucidate the mechanism of action of the effective agent and in turn the pathophysiology of the illness in that particular patient. Also these agents could be employed in potentiation studies to assess their applicability to larger populations. Working from the other direction, the ability to produce more specific diagnoses with homogeneous etiologies cannot come from large drug trials. Instead, the strategy of taking families where inheritance patterns suggest a strong genetic link to the development of affective disorder and where numerous individuals are available to be studied would allow for laboratory abnormalities to be searched for in what presumably is a homogeneous disorder and potentially highly specific treatments to be applied.

In conclusion, the various areas of research into affective disorder treatment do not stand alone in a vacuum but link together to allow validation of findings which leads to progress and finally clinical applicability. The chain is only as strong as its weakest link but as methodologies and instrumentation improve, the weakest links become drug action and diagnostic heterogeneity. The cornerstone of much of our research into treatments for affective disorders remains the drug trial. This study design which is necessary for the evaluation of the efficacy of a new treatment is probably non-conducive to the elucidation of highly specific drugs, is non-conducive to the study of homogeneous populations of patients with affective illness, and tends to be by its very nature an episode

versus illness study. We cannot discard the search for broad spectrum treatments but at the same time we must realize that if affective disorders have multiple etiologies, then the more specific the drug, the less it will meet government standards and the less commercially attractive it may become. Thus we will need to pass through a rather difficult phase in the development of new affective disorder treatments, before hopefully emerging at the other end with a sophistication in the treatment of affective disorders that parallels any other branch of medicine dealing with illnesses with putative multiple etiologies.

FIGURE 1

CURRENT RESEARCH INTERESTS	New Treatments	Predictors of Treatment Response			Treatment Monitoring	Mechanism of Action
		Clinical	Laboratory	Pharmacological Challenge		
A. Depression	2nd Generation Anti-depressants Potentiation Strategies	Diagnosis Clinical Variable	Transmitter Metabolite Uptake Inhibition Enzyme Receptor Binding Psychophysio-logical measures	DST TRH-TSH Amphetamine Clonidine Antidepressant	Plasma levels Transmitter Uptake Inhi-bition Receptor Binding	Receptor change
B. Mania	2nd Generation Anti-manics Potentiation Strategies	Diagnosis Clinical Variable	As above Membrane Character-istics	Lithium TRH-TSH	Plasma levels Other Tissue Levels	Membrane Effect Antikindling Effects Receptor Change
C. Illness	Carbamazepine Potentiation Strategies	Diagnosis Clinical Variable	As above Membrane Character-istics	As above	Plasma levels Other Tissue Levels	Membrane Effect Antikindling Effects Receptor Change

FIGURE 2

CURRENT CLINICAL APPLICATIONS	New Treatments	Predictors of Treatment Response			Treatment Monitoring	Mechanism of Action
		Clinical	Laboratory	Pharmacological Challenge		
A. Depression	2nd Generation Anti-depressants Potentiation Strategies					
B. Mania	2nd Generation Anti-manics Potentiation Strategies				Plasma levels	
C. Illness	Carbamazepine Potentiation Strategies	Diagnosis			Plasma levels	

REFERENCES

Angst, J., Weis, P., Grof, P., Baastrup, P.C. and Schou, M. Lithium prophylaxis in recurrent affective disorders. Br J Psychiat, 1970, 116, 604-614.

Asberg, M., Traskman, L. and Thoren, P. 5-HIAA in the cerebrospinal fluid. A biochemical suicide predictor? Arch Gen Psychiat, 1976, 33, 1193-1197.

Asberg, M., Bertillson, L., Tuck, D., Cronholm, B. and Sjoqvist, F. Indoleamine metabolites in the cerebrospinal fluid of depressed patients before and during treatment with nortriptyline. Clin Parmacol Ther, 1973, 14, 277-286.

Baldessarini, R.J. Treatment of depression by altering monoamine metabolism: precursors and metabolic inhibitors. Psychopharmaco Bulletin, 1984, 20:2, 224-23.

Baldessarini, R.J. Status of psychotropic drug blood level assays and other biochemical measurements in clinical practice. Am J Psychiat, 1979, 136:9, 1177-1180.

Ballenger, J.C. and Post, R.M. Carbamazepine (Tegretol) in manic-depressive illness: a new treatment. Am J Psychiat, 1980, 137, 782-790.

Beckman, H. and Goodwin, F.K. Central norepinephrine metabolism and the prediction of antidepressant response to imipramine or amitriptyline. Studies with urinary MHPG in unipolar depressed patients. Arch Gen Psychiat, 1975, 32, 17-21.

Bielski, R.J. and Friedel, R.O. Prediction of tricyclic antidepressant response. Arch Gen Psychiat, 1976, 33, 1479-1489.

Blombery, P.A., Kopin, I.J., Gordon, I.K., Markey, S.P. and Ebert, M.H. Conversion of MHPG to vanillylmandelic acid. Arch Gen Psychiat, 1980, 37, 1095-1098.

Brown, W.A. Use of dexamethasone suppression test in depression. In: Neurobiology of Mood Disorders, 1984, 18, 290-295.

Brown, W.A., Haier, R.J. and Qualls, C.B. Dexamethasone suppression test identifies subtypes of depression which respond to different antidepressants. Lancet, 1980, 1, 928-929.

Buchsbaum, M., Goodwin, F.K. and Murphy, D.L. Average evoked response in affective disorders. Relationship to imipramine response, unpublished manuscript.

Bunney, W.E., Hartmann, E.L. and Mason, J.W. Study of a patient with 48-hour manic-depressive cycles. II. Strong positive correlation between endocrine factors and manic defense patterns. Arch Gen Psychiat, 1965, 12, 619-625.

Cade, J.F. Lithium salts in the treatment of psychotic excitement. Med J Aust, 1949, 36, 349.

Charney, D.S., Menskes, D.B. and Heninger, G.R. Receptor sensitivity and the mechanism of action of antidepressant treatment. Arch Gen Psychiat, 1981, 38, 1160-1174.

Chouinard, G., Young, S.N., Bradwejn, J. and Annable, L. Tryptophan in the treatment of depression and mania. Advances in Biological Psychiatry, 1983, 10, 47-66.

Chouinard, G., Young, S.N. and Annable, L. Antimanic effect of clonazepam. Biological Psychiat, 1983, 18:4, 451-464.

Chouinard, G. (personal communication).

Cooper, T.B. and Carroll, B.J. Monitoring lithium dose levels: estimation of lithium in blood and other body fluids. J Clin Psychopharmaco, 1981, 1:2, 53-57.

Davidson, J.R.T., McLeod, M.N., White, H. and Raft, D. Red blood cell catechol-o-methyltransferase and response to imipramine in unipolar depressive women. Am J Psychiat, 1976, 133, 952-955.

Davis, J.M. Efficacy of tranquilizing and antidepressant drugs. Arch Gen Psychiat, 1965, 13, 552.

Delay, J. and Deniker, P. Le traitment des psychoses par une méthode neurolytique dérivée de l'hibernotherapie. In: Congrès des Médecins Aliénistes et Neurologistes de France, Luxembourg, 1952, pp. 497-502.

De Montigny, C., Cowinoyer, G., Morissette, R., Langlois, R. and Caille G. Lithium carbonate addiction in tricyclic antidepressant-resistant unipolar depression: correlations with the neurobiologic actions of tricyclic antidepressant drugs and lithium ion on the serotonin system. Arch Gen Psychiat, 1983, 40, 1327-34.

Dorus, E., Cox, N.J., Gibbons, R.D., Shaughnessy, R., Ghanshyam, N.P. and Cloninger, C.R. Lithium ion transport and affective disorders within families of bipolar patients. Arch Gen Psychiat, 1983, 40, 545-552.

Emrich, H.M., von Zerssen, D., Kissling, W., Moller, H-J. and Windorfer, A. Effect of sodium valproate in mania. The GABA-hypothesis of affective disorders. Arch Psychiatr Nervenkr, 1980, 229, 1-16.

Extein, I., Pottash, A.L.C., Gold, M.S. and Cowdry, R.W. Changes in TSH response to TRH in affective illness. In: Neurobiology of Mood Disorders, 1984, 19, 297-307.

Fawcett, J. and Siomopoulos, V. Dextroamphetamine response as a possible predictor of improvement with tricyclic therapy in depression. Arch Gen Psychiat, 1971, 25, 247-255.

Feighner, J.P. Clinical efficacy of the newer antidepressants. J of Clin Psychopharmaco, 1981, 1, 23-26.

Giannini, A.J., Houser, W.L., Loiselle, R.H., Giannini, M.C. and Price, W.A. Antimanic effects of verapamil. Am J Psychiat, 1984, 141:12, 1602-1603.

Glassman, A.H. and Roose, S.P. Delusional depression--a distinct clinical entity. Arch Gen Psychiat, 1981, 38, 424-427.

Goodwin, F.K., Prange, A.J. and Post, R.M. Potentiation of antidepressant effects by L-triiodothyronine in tricyclin nonresponders. Am J Psychiat, 1982, 139, 34-38.

Heninger, G.R., Charney, D.S. and Sternberg, D.E. Lithium carbonate augmentation of antidepressant treatment: an effective prescription for treatment-refractory depression. Arch Gen Psychiat. 1983, 40, 1335-42.

Hofmann von, G. and Bolen H. Therapeutic investigation of chronic manic depressive illness. Arzneimittelforsch, 1969, 19, 455-456.

Hollister, L.E., Davis, K.L. and Berger, P.A. Subtypes of depression based on excretion of MHPG and response to nortriptyline. Arch Gen Psychiat, 1980, 37, 1107-1110.

Hrdina, P.D. Antidepressant binding: implications for the mode of action and the biology of depression. Prog Neuro-psychopharmacol & Biol Psychiat, 1983, 7, 457-462.

Hullin, R.P. Minimum serum lithium levels for effective prophylaxis. In: Handbook of Lithium Therapy. England MTP Press, 1980.

Hwu, H.G., Rudorfer, M.V. and Clayton, P.J. Dexamethasone suppression test and subtypes of depression (Letter to Editor) Arch Gen Psychiat, 1981, 38, 363.

Insel, T.R. and Murphy, D.L. Pharmacologic response and subgroups of patients with affective illness. In: Neurology of Mood Disorders, 1984, 54, 854-862.

Jimerson, D.C. and Post, R.M. Psychomotor stimulants and dopamine agonists in depression. Neurology of Mood Disorders, 1984, 39, 619-625.

Kuhn, R. The treatment of depressive states with G#22355 (imipramine hydrochloride). Am J Psychiat, 1958, 115, 459.

Kukopulos, A., Reginaldi, D. Laddonada, P., Floris, G., Serra, G. and Tondo, L. Course of the manic-depressive cycle and changes caused by treatment. Pharmakopsychiatry, 1980, 13, 156-167.

Kupfer, D.J., Foster, F.G., Reich, L., Thompson, K.S. and Weiss, B. EEG Sleep changes as predictors in depression. Am J Psychiat, 1976, 133, 622-626.

Langer, G., Heinze, G., Reim, B. and Matussek, N. Reduced growth hormone responses to amphetamine in "endogenous" depressive patients: studies in normal, "reactive" and "endogenous" depressive schizophrenic, and chronic alcoholic subjects. Arch Gen Psychiat, 1976, 33, 1471-1475.

Lingjaerde, O. Platelet serotonin uptake inhibition as a basis for monitoring antidepressant drug treatment. J of Clin Psychopharmaco, 1984, 4:2, 76-81.

Louie, A.K. and Meltzer, H.Y. Lithium potentiation of antidepressant treatment. J Clin Psychopharmacol, 1984, 4:6, 316-320.

Maas, J.W., Fawcett, J.A. and Dekirmenjian, H. Catecholamine metabolism, depressive illness, and drug response. Arch Gen Psychiat, 1972, 26, 252-262.

Matussek, N., Ackenheil, M., Hippius, H., Muller, F., Schroder, H-Th., Schultes, H. and Wasilewki, B. Effect of clonidine on growth hormone release in psychiatric patients and controls. Psychiatr Res, 1980, 2, 25-36.

Okuma, T., Inanaga, K., Otsuki, S., Sorai, K., Takahishi, B., Hazama, H., Mori, A. and Watanabe, M. Comparison of the antimanic efficacy of carbamazepine and chlorpromazine: a double-blind study. Psychopharmacology, 1979, 66, 211-217.

Oppenheim, G. Estrogen in the Treatment of Depression: Neuropharmacological Mechanisms. Biological Psychiat, 1983, 18:6, 721-724.

Pandey, G.H., Dorus, E. and Davis, J.M. Lithium transport in human red blood cells: genetic and clinical aspects. Arch Gen Psychiat, 1979, 36, 902-908.

Paykel, E.S., Parker, R.R., Panrose, R.J.J. and Rassaby, E.R. Depressive classification and prediction of response to phenelzine. Br J

Post, R.M., Ballenger, J.C., Uhde, T.W. and Bunney, W.E. Efficacy of carbamazepine in manic-depressive illness: implications for underlying mechanisms. Neurobiology of Mood Disorders, 1984, 50, 777-811.

Post, R.M., Rubinow, D.R. and Ballenger, J.C. Conditioning, sensitization and kindling: implications for the course of affective illness. In: Neurobiology of Mood Disorders, 1984, 28, 432-459.

Potter, W.Z. and Linnoila, Markku. Tricyclic antidepressant concentrations: clinical and research implications. In: Neurobiology of Mood Disorders, 1984, 45, 698-705.

Prange, A.J. Estrogen may well affect response to antidepressant. J Am Med Assoc, 1972, 219, 143-144.

Prien, R.F., Caffey, E.M. and Klet, C.J. Factors associated with treatment success in lithium carbonate prophylaxis. Arch Gen Psychiat, 1974, 31, 189-192.

Quitkin, F.M., Rifkin, A., Klein, D.F. Monoamine oxidase inhibitors: A review of clinical effectiveness. Arch Gen Psychiat, 1979, 36, 749-759.

Quitkin, F.M., Kane, J., Rifkin, A., Ramos-Lorenzi, J.R. and Nayak, D.V. Prophylactic lithium carbonate with and without imipramine for bipolar 1 patients. Arch Gen Psychiat, 1981, 38, 902-907.

Raskin, A. Age-sex differences in response to antidepressant drugs. J Nerv Ment Dis, 1975, 159, 120-130.

Raskin, A. and Crook, T.H. The endogenous-neurotic distinction as a predictor of response to antidepressant drugs. Psychol Med, 1976, 6, 59-70.

Schildkraut, J.J. Norepinephrine metabolites as biochemical criteria for classifying depressive disorders and predicting responses to treatment and preliminary findings. Am J Psychiat, 1973, 130, 695-698.

Schou, M., Baastrup, P.C., Grof, P., Weis, P. and Angst, J. Pharmacological and clinical problems of lithium prophylaxis. Br J Psychiat, 1970, 116, 615-619.

Schwarcz, G., Halaris, A., Baxter, L., Escobar, J., Thompson, M. and Young, M. Normal thyroid function in desipramine nonresponders converted to responders by the addition of L-triiodothyronine. Am J Psychiat, 1984, 141:12, 1614-1616.

Sims, A. Monitoring lithium dose levels: estimation of lithium in saliva. In: Handbook of lithium therapy. England, MTP Press, 1980.

Spiker, D.G., Edwards, D., Hanin, I., Neil, J.F. and Kupfer, D.J. Urinary MHPG and clinical response to amitriptyline in depressed patients. Am J Psychiat, 1980, 137, 1183-1187.

Spring, G., Schweid, D., Gray, C., Steinberg, J., Horwitz, M. and Bond, D. A comparison of lithium and chlorpromazine in the treatment of manic states (a double-blind study). Am J Psychiat, 1970, 126, 1306.

Stancer, H.C. and Persad, E. Treatment of Intractable Rapid-Cycling Manic-Depressive Disorder with Levothyroxine. Arch Gen Psychiat, 1982, 39, 311-312.

Traskman-Bendz, L., Asberg, M., Bertilsson, L. and Sjostrand, L. Monoamine metabolites in CSF and suicidal behavior. Arch Gen Psychiat, 1981, 38, 631-636.

Van Praag, H.M. In search of the mode of action of antidepressants 5-HTP/tyrosine mixtures in depressions. Neuropharmacology, 1983, 22(33), 433-440.

Van Praag, H.M. Significance of biochemical parameters in the diagnosis, treatment, and prevention of depressive disorders. Biol Psychiat, 1977, 12, 101-131.

Zeller, E.A., Barsky, J., Fouts, J.R., Kirchheimer, W.F. and Van Order, L.S. Influence of isonicotinic acid hydrazide (INH) and 1-isonicotinyl-2-isopropyl hydrazine (IIH) on bacterial and mammalian enzymes. Experientia, 1952, 8, 349.

Zubenko, G.S., Cohen, B.M., Lipinski, J.F. and Jonas, J.M. Clonidine in the treatment of mania and mixed bipolar disorder. Am J Psychiat, 1984, 141:12, 1617-1618.

THE DEMENTIAS

THE CLINICAL DIAGNOSIS OF DEMENTIA

H. MERSKEY, M.D.
The University of Western Ontario

ABSTRACT

Dementia probably affects 11% of the population over 65. Alzheimer's Disease alone is responsible for dementia in at least 50% of autopsied cases. Another 15-20% approximately have combined Alzheimer's Disease and Multi-Infarct Dementia, 15% may have pure Multi-Infarct Disease. Diagnosis during life shows success rates varying from 43-70%. The characterisation of Alzheimer's Disease itself is better, with some 79-86% of cases being correctly identified. More investigation of Magnetic Resonance Imaging is needed at the present time. The better understanding of Alzheimer's Disease requires longitudinal studies during life with quantitative post mortem examination.

Dementia is usually defined nowadays as a loss of intellectual function accompanied also by memory disorder and other changes like impaired judgement or altered abstract thinking. There may be a loss of higher cortical function, e.g., dysphasia, agnosia or personality change. These are the essential features of the definition in the Diagnostic and Statistical Manual of the American Psychiatric Association, Third Edition (DSM-III, 1980). The Glossary of Mental Disorders in the International Classification of Diseases (ICD-9, 1977) describes an organic psychosis as one in which there is impairment of orientation, memory, comprehension, calculation, learning capacity and judgement. It then proceeds to define dementia as an organic psychosis in which impairment of mental function has developed to a degree that interferes grossly with insight, ability to meet some ordinary demands of life or to maintain adequate contact with reality. It is of a chronic or progressive nature, and if untreated, is usually irreversible and terminal. Both systems imply a threshold consideration so that some loss of intellectual function does not necessarily imply dementia. They specify various criteria or functions which may be lost and this is in accordance with the fact that many authors emphasize the global or generalized nature of the disturbance in dementia. Although I think it would be logical to treat any loss of intellectual function for organic reasons as a partial or focal dementia, this is not the conventional view nor the majority practice. Such focal losses are described more specifically as dysphasias or agnosias or as the amnestic syndrome in DSM-III. The ICD-9 incidentally, does not specifically separate an amnestic syndrome except under the different titles of retrograde amnesia (780.9) or Korsakov's syndrome, which may be alcoholic or non-alcoholic.

Dementia of course is often associated with delirium or clouding of consciousness and can only be soundly diagnosed when the effects of clouding of consciousness have been excluded or delimited. The notion of dementia also relies upon an organic cause which is either identified or presumed to be present. Otherwise, we talk of pseudo-dementia (Kiloh, 1961). Usually we expect pseudo-dementia to be due to depression, but classically dementia praecox was a schizophrenic illness and schizophrenia is sometimes a cause of pseudo-dementia in the elderly (Fisman, 1982).

It used to be common practice to describe dementia as a progressive and irreversible disorder. ICD-9 still describes it as usually chronic or progressive and irreversible unless treated. Sometimes, foolishly, it was made a criterion of dementia that it should be at least irreversible and often progressive. It was plainly mistaken to do that since dementias from head injuries, anoxia, anaesthetic mishaps, carbon monoxide poisoning, treated general paraesis of the insane, and other damaging events which usually did not recur were all well recognized. The suggestion of irreversibility was also not logical since there is no necessary reason why an organic intellectual deficit might not be recoverable. The successful treatment of some cases of hypothyroid dementia, B12 deficiency or normal pressure hydrocephalus effectively demolish this view that irreversibility is an essential feature of dementia.

Dementia affects large numbers of elderly people. In Britain as many as 11% over the age of 65 are believed to have dementia and in them perhaps 5.6% have a severe dementia (Kay et al., 1964, 1970). In Scandinavia figures as high as almost 16% (Essen - Möller, 1956) have been recorded. There are no comparable Canadian figures based upon a population survey. However it is reasonable to suppose that the situation in Canada is similar. The people are much the same as in those other countries and so are the common diseases. Routine clinical experience also suggests similar frequencies and patterns of illness in the elderly. The most important known difference is that the Canadian population is relatively young compared with that on the other side of the Atlantic. Thus, in Canada in 1981, 9.7% of the population were over 65. The world figure in 1980 was 5.8% over the age of 65, whilst the proportions in France in 1982 were 13.5% and in England and Wales in 1980, 15.1%. The older and more developed countries have the more aging populations (Fact Book on Aging in Canada, 1983).

Causes

The list of possible causes of dementia is endless and certainly runs to well over 100. Many of them are very rare. Common items include Alzheimer's Disease which accounts for about half of all cases or more on its own, multi-infarct disease which probably only accounts for about 10-15% of cases on its own and perhaps another 10-20% in association with Alzheimer's Disease. Cerebral tumour may account for as many as 5% of cases and alcoholism accounts for about 6% in some series (see Wells, 1977, page 250). Much more rarely, conditions occur like myxoedema, vitamin B12 deficiency, leukemic infiltrates, subcortical gliosis, Binswanger's Disease

and hippocampal sclerosis. Of course trauma, encephalitis, Huntington's Chorea, and Jakob-Creutzfeld Disease provide their quota and we can add Hallervorden-Spatz Disease, at least from our own series, even though it is largely thought to be a paediatric degenerative disorder. Normal pressure hydrocephalus is one treatable form of dementia but less frequent than the diagnosis. In patients who die with a diagnosis of dementia, the most common ultimate post mortem finding is that of Alzheimer's Disease whether presenile in origin or senile.

It used to be thought that Alzheimer's Disease and so-called senile dementia were distinct entities. The former was well characterized originally by early onset, senile plaques, neurofibrillary tangles and cerebral atrophy, with changes perhaps being most marked in the hippocampus. Gradually, the view developed that the two conditions had more features in common than differences and the two diagnoses were united (see Wolstenholme and O'Connor, 1970; Muller and Schwartz, 1978). Hence, it is now customary to talk of Alzheimer's Disease: senile type (ADST) and Alzheimer's Disease: presenile type (ADPST), or else of presenile dementia Alzheimer type (PSDAT) and senile dementia Alzheimer type (SDAT). Despite the unification of the overall concept the suggestion has now begun to emerge that SDAT may be split into two groups that are different, at least chemically and pathologically. Rossor et al. (1984) describe neurochemical differences of patients dying in their seventh and eighth decades compared with those dying in their ninth and tenth decades. The younger patients had a widespread and severe cholinergic deficit together with abnormalities of noradrenaline, gamma-amino-butyric-acid and somatostatin whereas the older ones had a relatively pure cholinergic deficit confined to the temporal lobe and hippocampus. Thus, it is now suggested that there may be some important differences between subtypes of Alzheimer's Disease.

I indicated just now that the commonest cause of dementia was Alzheimer's Disease. Dr. Ball will discuss the pathological criteria for Alzheimer's Disease but to date the best evidence on its frequency comes from pathological reports of demented patients examined post mortem. There are at least nine reports (Peters & Struck, 1959; Rothschild, 1941; Newton, 1948; Raskin and Ehrenberg, 1956; Delay and Brion, 1962; Simon and Malamud, 1965; Lauter and Meyer, 1968; Jellinger, 1976; and Tomlinson, 1977) which provide figures for the relative frequency of multi-infarct dementia and Alzheimer's Disease in dying patients suffering from functional mental illness or dementia. The first seven reports describe "senile dementia" in between 7.5% to 53% of psychiatric hospital autopsies. "Arteriosclerosis" accounted for from 23% to 92.5% of cases. Peters and Struck (1959) described no mixed cases but the others found from 13.6% to 54.8% to be examples of both types of dementia. In 31 cases dying with dementia Delay and Brion (1962) found 17 mixed cases. Lauter and Meyer (1968) reporting autopsies by M. Sabuncu on 35 patients found about one-third with "arteriosclerosis", one-third with senile dementia and 40% with mixed pathology. In 50 cases, Tomlinson (1977) observed that, depending on the criteria used for multi-infarct disease and Alzheimer's Disease, some 50-56% of hospitalized patients with dementia had pure Alzheimer's Disease, 16-18% had multi-infarct dementia and 12-18% had both conditions. The remainder had dementia for a variety of causes. Tomlinson's figures have been the most

generally accepted and suggest a remarkable overlap, incidentally, between Alzheimer's Disease and multi-infarct disease, which is usually explained as being due to two illnesses causing dementia more readily than one alone.

The post mortem studies depended largely upon institutional populations. However a number of clinical studies support the same conclusions in a general fashion. I referred to these earlier. Amongst these studies Rabins (1981) suggests that 10% of patients are treatable. Marsden and Harrison (1972) found that about 21% were treatable; Kiloh and Smith (1981) found 17% treatable, Delaney (1982) 23% treatable; and at hospitals of initial referral Harrison and Marsden (1977) and Freemon (1981) found that more than one-third of patients first seen are treatable.

It is obviously important to diagnose the minority of patients with treatable organic conditions. Besides that the group with depression requires careful separation from the remainder with organic disease. This group of patients with depression or other functional illness may be recoverable and there is active current interest in them (Bienenfeld and Hartford, 1982; Fisman, 1982; Janowsky, 1982; Himmelhoch et al. 1982; Reifler, 1982). One qualification should be expressed here. Dementia may occur with a superimposed pseudo-dementia as Reifler (1982) has noted. Perhaps, too, pseudo-dementia may be a precursor of dementia with a bad prognosis. Kral (1982) followed 22 patients with pseudo-dementia for four to 18 years and found that 11 had died suffering from dementia and nine more developed it. Only two survived without evidence of dementia. Nevertheless, the treatment of depression itself is obviously valuable even if dementia will supervene later.

It becomes further important to diagnose AD if an effective method of treatment should be available. I therefore propose to consider next the reliability of the diagnosis of dementia in general and Alzheimer's Disease in particular.

If one looks for a completely accurate ante-mortem diagnosis in all patients the figures are not encouraging when they are compared with the post mortem findings. Constantinidis (1978) reports 43% accuracy; Muller and Schwartz (1978) 55%; Nott and Fleminger (1975) 70%; Ron et al. (1980) 70% on clinical follow-up, and in the first 46 patients of our own series we achieved 63% accuracy.

The attempt to diagnose the specific disorder of Alzheimer's Disease is more successful. McLachlan (1984) reports 86% in an unpublished series and Sulkava et al. (1983) carefully defined a group of patients and then achieved 82% accuracy in diagnosis. Our own results appear in Table I. Complete consonance between four major diagnostic categories was obtained in 63%. Correct prediction of the presence of Alzheimer's Disease was achieved in 79% of cases and of multi-infarct dementia in 64%. These results compare with the best in the literature and are very disappointing. They were based upon the following procedure for the diagnosis of AD and MID. First a careful consideration of the detailed history, and a physical examination by a neurologist; second, the exclusion of all other detectable causes of dementia by the usual blood tests and with minimally

invasive routines; third, the review of EEG and CT scan findings; and fourth, serial psychometry and follow-up of the patients who were under continuing observation.

Refinements to the existing procedure, based on this experience, may help us in the future. One point which we have already determined is that the EEG which is frequently abnormal in AD (Gordon and Sim, 1967) is occasionally normal right through to within a few months of death, and certainly at advanced stages of the disease. We have at least four patients in whom this was the case out of the 46 who are considered here.

Both the EEG and CT scan have a general relationship to the severity of AD. The relationship is not enough for absolute diagnostic purposes, although a normal EEG in a case of moderately severe dementia will tend to raise a question since a normal EEG is admittedly rare in advanced AD. However, the quantitative assessment of the EEG could conceivably serve as a means of evaluation of change for better or worse and the CT scan may help to give some measure of the patient's condition albeit only in broad categories.

The CT scan is probably more useful in the recognition of infarction. Radiologists can readily observe areas of lucency in the scan, some of which are correctly identified with confidence as infarcts. Others do not appear to correspond to clinical or pathological evidence of infarction and their interpretation is a matter for current enquiry. The neurological examinations which we have relied upon were mostly supervised by Dr. Hachinski. In reaching the diagnoses described above we have naturally relied upon the Ischaemic Scale Score (Hachinski et al., 1974). Hachinski is the first to point out that the score has not been properly validated. Harrison et al. (1977) undertook a partial validation based upon neuroradiological findings and Rosen et al. (1980) published a small series of cases on the basis of which they recommended a modified scale score. It is not clear to us that the modified scale has any advantages yet over the larger one. It will clearly be important to use our own material to determine the best elements in the scale by comparison with post mortem findings.

Imaging Techniques

Several more technical methods are now available for the investigation of the metabolic function of the brain or its anatomical structure. It has been known for a number of years that cerebral blood flow is reduced with Alzheimer's Disease but the findings are too general and non-specific to be of much help. They were well reviewed as long ago as 1978 (Ingvar et al., 1978) and little seems to have been put forward about this approach since then. The CT scan has already been mentioned. We have hopes to produce more refined measures based upon the CT scan and there is still a need to look at some aspects of it: for example, better true volumetric assessments of the size of the ventricles are desirable and thin slice examination of the hippocampus is a technique which we have in process at the moment. However, at the present time it has to be said that the main function of the CT scan is to determine the presence or absence of infarcts

and to exclude other disease. It also bears a general relationship to the extent of the disease although this is only of limited use in quantification. The more recent techniques of neuroimaging are likely to be dealt with by Dr. McGeer. I have no experience of PET scanning but members of our group are now working on the initial phase of investigations with Magnetic Resonance Imaging. It is possible that that procedure may have no more fine resolution than CT scan but my radiological colleagues have at least some data which indicates a need for further examination of the possible differences in T1 and T2 relaxation times in MRI investigations. This evidence so far is based only upon post mortem preparations. A large increase in T2 values relative to control slides was obtained from the brains of patients with Alzheimer's Disease which contained plaques and tangles (Karlik, 1985).

Clinicopathological Definition of Subgroups

Dr. Ball will be describing the criteria by which Alzheimer's Disease may be defined post mortem neuropathologically on a routine clinical basis, and also on a more exacting quantitative level. When firm quantitative criteria are applied, not all cases who have Alzheimer's Disease post mortem on routine examination also meet the full quantitative criteria. We are currently looking into some of the differences between these groups. Dr. Ball will present data on 37 brains which have been fully examined. Only 21 of those brains meet the most rigorous morphometric standards for Alzheimer's Disease. Another 16 fall short of those standards which Dr. Ball will detail but did not have any other diagnosis. I present here, data on a lesser number of patients which has been compiled in regard to the clinical condition of the individuals during life. The initial findings on 16 of the patients with full criteria, called "typical" Alzheimer's Disease, and eight with incomplete criteria, called "atypical" AD, are shown in Table 2. It appears that the "atypical" patients are older at death than the typical ones, but have been ill for the same length of time. Their illnesses seem to be less extensive pathologically. This is perhaps borne out by the fact that the choline acetyltransferase levels (which are not shown) are higher in the older patients, although choline acetyltransferase normally declines with age as well as with Alzheimer's Disease.

Further, although it affects both typical and atypical groups, there is one other finding which has a relationship to age. In the patients who died, there were four who had normal EEG's at periods of 1, 4, 22 and 31 months from death, when their psychological test scores were moderately decreased in two cases and severely decreased in two more. These patients ranged in age at death from 79 to 86 years, provisionally suggesting that the well recognized EEG abnormality of AD is only found consistently in younger patients. This agrees well with Gordon and Sim's (1967) report since their patients who all had abnormal EEG's had presenile dementia. It seems overall that our results suggest that age has important effects upon the pathological process which, perhaps surprisingly, may be milder in the older cases and this is in accordance with the work of Rossor et al. (1984).

To revert to the main question, there now seem to be two principal themes in the diagnosis of AD. One is to discriminate some subgroups, when possible. The other is to find new and more promising techniques to support and explore during life. Magnetic Resonance Imaging is probably the most popular topic here but is only just beginning to be explored.

Acknowledgements

This work has been supported by grants to the U.W.O. Dementia Study Group with Dr. M.J. Ball as Principal Investigator as follows: Program Grant PG/21; Medical Research Council of Canada; Grant #ROI AGNS 03047 U.S. National Institute on Aging; and also the Ontario Mental Health Foundation. I thank all the members of the U.W.O. Dementia Study Group for their participation in the work reported.

TABLE 1

NEUROPATHOLOGICAL DIAGNOSIS AND CLINICAL DIAGNOSIS

Clinical Diagnosis	Alzheimer's Disease	Mixed A.D. & M.I.D.	Multi-Infarct Disease	Other
A.D.	18	0	0	2
Mixed A.D. & M.I.D.	4	5	1	1
M.I.D.	2	0	1	0
Other	3	2	2	5
Total	27	7	4	8

34 11

Exact agreement 29/46 = 63%

Recognition of A.D. 27/34 = 79%

Recognition of M.I.D. 7/11 = 64%

TABLE 2

INITIAL FINDINGS IN TWO GROUPS WITH ALZHEIMER'S DISEASE

	"TYPICAL"	"ATYPICAL"
Age at onset	65.6 years (N = 15)	71.43 years (N = 7)
Age at death	73.75 years (N = 16)	80.13 years (N = 8)
Duration	8.8 years (N = 15)	8.6 years (N = 7)
Number with normal final EEG/abnormal final EEG	1/16	3/8

Age at onset, $t = 1.71$, 20df, $P<0.20$ N.S.

Age at death, $t = 2.30$, 22df, $P<0.05$

REFERENCES

Bienenfeld, D. and Hartford, J.T. Pseudodementia in an elderly woman with schizophrenia. Am J Psychiat, 1982, 139, 1:114-115.

Constantinidis, J., Richard, J. and de Ajuriaguerra, J. Dementias with senile plaques and neurofibrillary changes. In: Studies in Geriatric Psychiatry. Ed. A. Isaacs and F. Post. Wiley: Toronto, 1978.

Delaney, P. Dementia: The search for treatable causes. Southern Med J, 1982, 75, 6:707-709.

Delay, J. and Brion, S. Les démences tardives. Ed. Mason & Cie: Paris,1962, pp. 180-230.

Diagnostic and Statistical Manual. Third Edition (DSM-III). Am Psychiat Assoc, 1980.

Essen-Möller, E. Individual traits and morbidity in a Swedish rural population. Acta Psychiat Neurol Scand Suppl, 1956, 100, 1-160.

Fisman, M. Pseudodementia and schizophrenia. Am J Psychiat, 1982, 139, 6:846.

Freemon, F.R. Organic Mental Disease. Spectrum Publications: New York, 1981.

Gordon, E.B. and Sim, M. The EEG in presenile dementia. J Neurol Neurosurg and Psychiat, 1967, 30:285-291.

Hachinski, V.C., Lassen, N.A. and Marshall, J. Multi-infarct dementia: A cause of mental deterioration in the elderly. Lancet, 1974, 2:207-209.

Harrison, M.J.G. and Marsden, C.D. Progressive intellectual deterioration. Arch Neuro, 1977, 34:199.

Himmelhoch, J.M., Auchenbach, R. and Fuchs, C.Z. The dilemma of depression in the elderly. J Clin Psychiat, 1982, 43, 9:26-32.

Ingvar, D.H., Brun, A., Hagberg, B. and Gustafson, L. Regional cerebral blood flow in the dominant hemisphere in confirmed cases of Alzheimer's Disease, Pick's Disease, and multi-infarct dementia: Relationship to clinical symptomatology and neuropathological findings. In: Alzheimer's Disease: Senile Dementia and Related Disorders, (Aging, Vol. 7). Ed. R. Katzman, R.D. Terry and K.L. Bick. Raven Press: New York, 1978, pp. 203-211.

Janowsky, D.S. Pseudodementia in the elderly: Differential diagnosis and treatment. J Clin Psychiat, 1982, 43, 9:19-25.

Jellinger, K. Neuropathological aspects of dementias resulting from abnormal blood and cerebrospinal fluid dynamics. Acta Neurologica (Belgium), 1976, 76:83-102.

Karlik, S. Unpublished data, 1985.

Kay, D.W.K., Beamish, P. and Roth, M. Old age mental disorders in Newcastle-upon-Tyne. I. A study of prevalence. Br J Psychiat, 1964, 110, 146-158.

Kay, D.W.K., Bergmann, K., Foster, E.M., McKechnie, A.A. and Roth, M. Mental Illness and hospital usage in the elderly: A random sample followed up. Compre Psychiat, 1970, 11:26-35.

Kiloh, L.G. Pseudo-dementia. Acta Psychiatrica Scan, 1961, 37, 336-351.

Kral, V.A. Depressive pseudodemenez und senile demenz vom Alzheimer Typ. Nervenartz, 1982, 53:284-286.

Lauter, H. and Meyer, J.E. Clinical and nosological concepts of senile dementia. In: Senile Dementia: Clinical and Therapeutic Aspects. Ed. Ch. Muller and L. Ciompi. Hans Huber Publishers: Bern and Stuttgart, 1968, pp. 13-26.

Marsden, C.D. and Harrison, M.J.G. Outcome of investigation of patients with presenile dementia. Br Med J, 1972, ii:249-252.

McLachlan, D. Unpublished observations: Personal communication, 1984.

Muller, H.F. and Schwartz, G. Electroencephalograms and autopsy findings in geropsychiatry. J Gerontol, 1978, 33, 4:504-513.

Newton, R.D. The identity of Alzheimer's Disease and senile dementia and their relationship to senility. J Ment Sci, 1948, 94:225-249.

Nott, P.N. and Fleminger, J.J. Presenile dementia: The difficulties of early diagnosis. Acta Psychiat Scan, 1975, 51:210-217.

Peters, G. and Struck, G. Pathomorphologische Befunde bei chronish verlaufenen hirnatrophischen Prozessen. Fortschr Neurology, Psychiat, 1959, 27:549.

Rabins, P.V. The prevalence of reversible dementia in a psychiatric hospital. Hosp and Commun Psychiat, 1981, 32, 7:490-492.

Raskin, N. and Ehrenberg, N. Senescence, senility and Alzheimer's Disease. Am J Psychiat, 1956, 98:324.

Reifler, B.V. Arguments for abandoning the term pseudodementia. J Am Geriat Soc, 1982, 30, 10:665-668.

Ron, M.A., Toone, B.K., Gerralda, M.E. and Lishman, W.A. Diagnostic accuracy in presenile dementia. Br J Psychiat, 1979, 134:161-168.

Rosen, W.G., Terry, R.D., Fuld, P.A., Katzman, R. and Peck, A. Pathological verification of ischemic score in differentiation of dementia. Ann Neurol, 1980, 7:486-488.

Rossor, M.N., Iversen, L.L., Reynolds, G.P., Mountjoy, C.Q. and Roth, M. Neurochemical characteristics of early and late on-set types of Alzheimer's Disease. Br Med J, 1984, 228:961- 964.

Rothschild, D. The clinical differentiation of senile and arteriosclerotic psychoses. Am J Psychiat, 1941, 98:324-333.

Simon, A. and Malamud, N. Comparison of clinical and neuropathological findings in geriatric mental illness. In: Psychiatric Disorders in the Aged. Geigy Ltd.: Manchester, 1965, pp. 322.

Smith, J.S. and Kiloh, L.G. The investigation of dementia: Results in 200 consecutive admissions. Lancet, 1981, 824-827.

Sulkava, R., Haltia, M., Paetau, A., Wikstrom, J. and Palo, J. Accuracy of clinical diagnosis in primary degenerative dementia: Correlation with neuropathological findings. J Neurol Neurosurg and Psychiat, 1983, 46:9-13.

The Glossary of Mental Disorders. International Classification of Diseases, 1979. World Health Organization, Geneva.

Tomlinson, B.E. Morphological changes and dementia in old age. In: Aging and Dementia. Ed. W.L. Smith and M. Kinsbourne. Spectrum Publications: New York, 1977, pp. 25-26.

Wells, C.E. and Buchanon, D.C. The clinical use of psychological testing in evaluation of dementia. In: Dementia. 2nd Edition. Ed. C.E. Wells. F.A. Davis: Philadelphia, 1977.

Wolstenholme, G.E.W. and O'Connor, M. Alzheimer's Disease and Related Conditions. Churchill: London, 1970.

RECENT ADVANCES IN THE PATHOLOGICAL DIAGNOSIS OF SENILE DEMENTIA ALZHEIMER TYPE

M.J. BALL, M.D.
Department of Pathology
University of Western Ontario

Introduction

For the last couple of years, the Dementia Study Group at the University of Western Ontario has been organizing a prospective, clinico-pathological investigation of organic dementia of the Alzheimer type. A large cohort of organically demented persons suffering either from the presenile or the senile type of onset of this commonest form of dementia is being investigated by systematic neurological and psychiatric examinations, specialized psychometric instruments, quantitative 16-channel electroencephalography, and quantitative computerized tomographic radiography. Clinically these patients are categorized as having pure or uncomplicated Alzheimer's Disease; multi-infarct dementia; "mixed" dementia (a combination of the first two categories); or "other types." The London Psychogeriatric Rating Scale, a 36-item questionnaire developed in our institutions and showing high inter-rater reliability (Hersch et al., 1978a) as well as clinical validity (Hersch et al., 1978B, 1980) is carried out as a behavioural observation scoring test every three months. The Extended Scale for Dementia, which is applied to patients intellectually at the level of or below the lower end of the Wechsler Adult Intelligence Scale, is utilized to test cognitive function every six months for those patients in the earlier stages of their illness. Both the LPRS and ESD scores correlate closely with EEG and CT scan measures in patients with Alzheimer's Disease (Merskey et al., 1980). The electroencephalogram is repeated annually, and the tomographic scan every second year.

A control group of cognitively intact normal subjects, matched for age, has volunteered to undergo an identical series of clinical tests at our medical school. Careful liaison with the nearest relative of the demented individuals has provided us with a compliance rate for autopsy permission exceeding 75%. Because of the proximity of the regional nursing homes to University Hospital, fresh brain tissues become available both for neurohistological and for biochemical studies usually within 10-12 hours post mortem, and frequently as early as 2-4 hours.

Quantitative histomorphometry as well as cytoarchitectonic topography are carried out on five different lesions of the grey matter -- neurofibrillary tangles of Alzheimer, "senile" or neuritic plaques, granulovacuolar degeneration of Simchowicz, rod-like bodies of Hirano, and nerve cell loss. These specialized very labour-intensive techniques, which utilize a semi-automated sampling stage microscope (Wild M501), have been developed in my laboratory in order that neuronal lesions may be accurately quantified according to stereological principles, and mapped on anatomical "scattergrams" with the aid of a potentiometrically linked XY pen

recorder. From multiple samples both of the same neocortical and limbic regions used in our neuropathological morphometry, biochemical assays are also being performed of markers for several neurotransmitter systems--in particular, the activity of choline acetyltransferase as a reflection of the status of the cholinergic family of neurons. Brains obtained from our control group provide normative data.

Because of the availability of fresh CNS tissue harvested at post-mortem in this fashion, Drs. R. Jane Rylett and E. Howard Colhoun have been able to perform unique studies of transport uptake mechanisms of choline, utilizing synaptosomal preparations prepared from such samples (Rylett et al., 1983).

Pathological Criteria

The neuropathological criteria which might be accepted as some uniform standard for the diagnosis of Alzheimer's Disease remain an important goal for pathologists studying the organic dementias. At a Research Workshop on the Diagnosis of Alzheimer's Disease recently sponsored by the National Institutes of Health, the report of the Panel on Neuropathology (which I was privileged to chair) provided some minimum microscopic criteria necessary for a community hospital pathologist to establish a histological diagnosis of Alzheimer's Disease, when presented with a brain from an autopsy in the absence of any informative clinical history (Research Workshop, 1983). In addition, however, our own Dementia Study Group considers it critically important to test more rigorous research criteria for this diagnosis, utilizing the prospective longitudinal human project in which clinical parameters of cognitive decline documented in vivo can then be compared with cerebral histological abnormalities in the same individuals after death.

Grossly, the brain of a patient with Alzheimer's Disease of the uncomplicated variety should have not more than a total of 50 c.c.'s of formalin-fixed tissue exhibiting any area of infarction, after all the one-centimeter thick coronal slices serially through the cerebral hemispheres have been surveyed. In addition, there should be no other gross or microscopic features of any additional cause for dementia, such as subdural hematoma, Pick inclusion bodies, or a primary cerebral neoplasm. From earlier data published in my laboratory (Ball, 1977), in which retrospective analysis of the hospital charts did not permit a quantitative psychometric evaluation of the degree of dementia, I have felt that in every case diagnosed clinically as uncomplicated Alzheimer's Disease the serially sectioned hippocampal formation should probably demonstrate:

(a) at least 20 tangle-containing, nucleolated nerve cells per cubic millimeter--the Adjusted Tangle Index;

and/or

(b) at least 55 nucleolated nerve cells per cubic millimeter exhibiting granulovacuolar degeneration -- the Adjusted Granulovacuolar Index;

and/or

(c) a neuronal population of less than 5,600 nucleolated neurons per cubic millimeter.

Recently, the entire hippocampal formation was analyzed by quantitative morphometry in 37 cases of Alzheimer's Disease. Twenty-one of these had been fully investigated during life and had demonstrated all the DSM-III (American Psychiatric Association) criteria for the clinical diagnosis of organic dementia Alzheimer type. No other dementing illness was present. Some investigations were incomplete on the other 16 because of their frail clinical condition, but otherwise they met the same criteria. In 15 of these 37 patients, the degree of histological deterioration of one of these three hippocampal lesions does not quite match the above morphometric limits, and in four of these 15 cases the severity of two of the types of lesions falls somewhat short. For investigative purposes, we have called these 15 brains "less typical" examples of Alzheimer's Disease.

We are, therefore, now comparing the various clinical parameters, including duration and slope of illness, radiological changes, and electroencephalographic abnormalities, with those others which seem more typically severe, in order possibly to refine further our pathological definitions of this disease.

Conclusion

This type of quantitative approach is currently in progress in my laboratory, and provides a golden opportunity for the morbid anatomist to amplify further the neuropathological definition of what constitutes Alzheimer's Disease (both senile dementia of the Alzheimer type, and presenile Alzheimer's Disease). These efforts will eventually permit us to sort out which of the many histological and biochemical abnormalities described in this condition most appropriately account for the decline of cognitive function in this irreversible degenerative disorder.

Until these kinds of crucial questions can be more fully answered, attempts to understand the pathogenetic mechanisms giving rise to such cellular abnormalities by using animal models will be fraught with the distinct possibility of studying reactive epiphenomena, instead of the primary cellular lesions.

Acknowledgements

This study has been supported in part by grants from the Medical Research Council of Canada (PG21), the National Institutes of Health (AGNS03047), and the Ontario Mental Health Foundation (858). Ms. Nancy

Joslin kindly prepared the manuscript. The author thanks all members of the University of Western Ontario Dementia Study Group for their generous collaborative assistance.

REFERENCES

Ball, M.J. Neuronal Loss, Neurofibrillary Tangles and Granulovacuolar Degeneration in the Hippocampus with Aging and Dementia: A Quantitative Study. Acta Neuropath, (Berl.) 1972, 77, 111-118.

Hersch, E.L., Csapo, K.G. and Palmer, R.B. Development of the London Psychogeriatric Rating Scale. London Psychiatric Hospital Bulletin, 1978a, 8, 1, 3.

Hersch, E.L., Kral, V.A. and Palmer, R.B. Clinical Value of the London Psychogeriatric Rating Scale. Journal of the American Geriatrics Society, 1978b, 8, 348.

Hersch, E.L., Merskey, H. and Palmer, R.B. Prediction of Discharge From a Psychogeriatric Unit. Canadian Journal of Psychiatry, 1980, 25, 234.

Merskey, H., Ball, M.J., Blume, W.T., Fox, A.J., Fox, H., Hersch, E.L. and Palmer, R.B. Relationship Between Psychological Measures and Cerebral Organic Changes in Alzheimer's Disease. Canadian Journal of Neurological Sciences, 1980, 7, 45.

Report of the Panel on Neuropathology, Research Workshop on the Diagnosis of Alzheimer's Disease, December 4-7, 1983, Bethesda, Maryland. Sponsored by National Institute on Aging, National Institute of Neurological and Communicative Disorders and Stroke, and National Institute of Mental Health (in press).

Rylett, R.J., Ball, M.J. and Colhoun, E.H. Evidence for High Affinity Choline Transport in Synaptosomes Prepared from Hippocampus and Neocortex of Patients with Alzheimer's Disease. Brain Research, 1983, 289, 169.

ALCOHOLISM

DIAGNOSIS AND TREATMENT OF ALCOHOLISM

C.S. MELLOR, M.D.
Health Sciences Centre

ABSTRACT

The criteria for the diagnosis of alcoholism as defined in DSM-III are considered and their validity reviewed in the light of other diagnostic criteria. Biological methods of diagnosing alcoholism using laboratory tests point to the usefulness of the GGT and MCV. Reference is made to the computerization of laboratory test results thus providing an automated diagnosis. The main biological advances in the area of physical dependence and its treatment are considered including the question of the opiate theory and TIQs, hypoxia, and endocrine changes particularly of vasopressin levels.

This paper provides a short review of the major biological developments in the field of alcoholism; with an emphasis upon recent developments in diagnosis and the Alcohol Withdrawal Syndrome (AWS).

Diagnosis

Biomedical research must build upon firm diagnostic foundations. There is no doubt that the DSM-III[1] now pervades all psychiatric research and will continue to do so.

The DSM-III has three major diagnostic groupings for alcoholism; alcohol abuse, alcohol dependence and the alcohol organic mental disorders. Alcohol abuse is defined as a pattern of pathological use for at least a month that causes impairment in social or occupational functioning. Alcohol dependence is defined as either pathological use of alcohol or impairment of social or occupational functioning due to alcohol, together with either tolerance or dependence. Somewhat surprisingly, because the definition of alcoholism is always a controversial subject, these criteria have not been subjected to the same critical reviews as those for other psychiatric diagnoses.

Nevertheless, the validity of the DSM-III criteria and their forerunners the Research Diagnostic Criteria and the Feighner criteria for diagnosing alcoholism have been questioned[2]. These authors point out, as have other critics of the DSM-III, that high levels of diagnostic reliability are not necessarily rewarded by validity. Their own investigations lead them to conclude that these criteria fail to identify a substantial proportion of clinic cases, are biased in the cases that are identified, and are not reliable through time.

Boyd et al[3] have made comparisons between the DSM-III and other criteria and have demonstrated varying levels of sensitivity in population studies. Consequently, they have developed the Alcohol Polydiagnostic Interview[4], which attempted to synthesize and operationalize the '10 leading definitions' of alcoholism used in the United States and Europe. A full evaluation of this instrument has not been published. The top 10 criteria are listed below.

TABLE I

List of 10 Diagnostic Criteria

1. DSM-III alcohol abuse
2. DSM-III alcohol dependence
3. Research Diagnostic Criteria
4. Feighner Criteria
5. National Council on Alcoholism
6. Jellinek's definitions
7. ICD-9
8. Present State Examination
9. Iowa Alcoholism Stages Index
10. Michigan Alcoholism Screening Test (MAST)

The Michigan Alcoholism Screening Test is perhaps the most widely used and has been the subject of further refinement[5].

Attempts have been made to circumvent the difficulties attendant upon what have been termed "Trouble-due-to-drinking scales" by using laboratory tests to make the diagnosis. These have been recently reviewed by Morgan[6], and the most frequently used are listed below.

TABLE II

Commonly Used Markers for the Detection of Alcoholism

Blood alcohol concentration
Erythrocyte δ-aminolevulic acid and dehydrase
Mean corpuscular volume (MCV)
Plasma α-lipoproteins
α-amino-n-butyric acid: leucine ration
Serum asparate aminotransferase (SGOT)
Serum γ-glutamyltransferase (GGT)
Serum transferrin

It was concluded from this review that the MCV, SGOT and GGT were most useful in identifying active alcoholism. Their diagnostic value is limited by a sensitivity about 60% for MCV and GGT combined and a similar level of specificity[7]. (The MCV is more likely to be raised in moderate

sustained than intermittent heavy drinkers.) Attempts have been made to improve the diagnostic value of laboratory investigations by using a battery of biochemical tests which are then subject to discriminant analysis. The process can be computerized to provide a rapid automated screening test for alcoholism and sensitivity and specificity rates of about 80% are claimed for this procedure[8].

Finally, the effectiveness of combining screening tests based upon rapid interviews and laboratory tests have been examined[9]. In this study glutamate dehydrogenase was the best laboratory test being superior to the GGT and being comparable to the short screening tests in efficacy. When it was combined with the Reich interview[10] a sensitivity of 100% was claimed.

The DSM-III diagnostic criteria for those organic conditions that are attributed to alcoholism do not present difficulty although one might quibble at the Alcohol Withdrawal Syndrome being regarded as a diagnostic entity quite separate from Delirium Tremens.

Biological Predisposition to Alcoholism

Familial aspects of persons with high risk for alcohol are considered in Dr. Dongier's paper. In addition, however, reference should be made to the question of laterality. There is a report of increased frequency of mixed laterality in alcoholics[11] and a reduction in alcoholics of the predominant left sided asymmetry observed in the CAT scans of normal Caucasians.[12] The question of whether this is cause or effect receives a partial answer from a large normal population study in which a higher frequency of drinking was associated with left-handedness[13].

The reduced risk of motion sickness in alcoholics recently reported[14] is of some interest because of problems encountered trying to induce conditioned taste aversions to alcoholic beverages in alcoholism using motion sickness as the aversive stimulus[15]. The question remains whether a resistance to vestibular stimulation by alcohol, as well as motion, predisposes to alcoholism, or is an acquired characteristic.

Alcohol Withdrawal Syndrome (AWS)

Any rational treatment of the withdrawal syndrome must rest upon an understanding of the pathogenesis of the condition. Some of the pathophysiological disturbances which might play a role in the development of this condition will be considered.

Magnesium: The resemblance between the clinical manifestations hypomagnesia and the AWS has long been remarked upon. The early findings that serum Mg++ was reduced in the AWS supported this possibility. The serum magnesium abnormalities certainly appear to parallel the course of delirium tremens[16]. However, the failure to find reduction of the CSF Mg++

in experimental withdrawal states[17] and the fact that most patients in AWS with low Mg++ recover without attention to their Mg++ suggest that it is a consequence or a complication of this disorder rather than a cause.

Acetaldehyde Condensates: A link between alcohol and opiate addiction was suggested first in 1970.[18] The acetaldehyde formed by the metabolism of alcohol would form, it was postulated, a new class of compounds, the tetraisoquinolones (TIQ's) by condensation with the various catecholamines. This theory has provoked considerable speculation. It has been suggested that the simpler TIQ's may serve as false neurotransmitters, competing at the cerebral synapses for catecholamines, thus increasing catecholamine synthesis and consequently an excessive, catecholamine induced, cerebral stimulation when alcohol is withdrawn. Another hypothesis is that more complex TIQ's which resemble opiate alkaloids have an opiate-like action and by their attachment to the opiate receptors produce addiction and an opiate withdrawal syndrome when alcohol is withdrawn. Despite extensive in vitro and animal studies the hypothesis is unproven.

Naloxone does not provoke the AWS, but has been shown to rapidly reverse alcohol induced coma. In one study this occurred in 20% of patients.[19] Alpha-adrenergic agonists such as clonidine and loxefidine have been reported to block opiate withdrawal syndromes[20,21]. This is attributed to the blockade of noradrenergic overactivity. A recent open study suggests that clonidine is also effective in alcohol withdrawal[22], but this has not been confirmed in recent blind studies using rats[23].

Hypoxia: A reduction in cerebral oxygenation has been cited by several workers as a feature of AWS[24], and other studies have pointed to a concomitant reduction in the cerebral blood flow[25]. This work has recently been reviewed by Gillman and Lichtigfeld[26] who claim that pure oxygen will clear approximately two thirds of the symptoms of AWS. In addition when this is combined with nitrous oxide, an opiate agonist, there is further improvement.

Na+, K+ and Ca++: The significance of the role of disturbed electrolyte balance in AWS is not known. It does appear that increased Ca++ sensitivity in the presynaptic terminal may play a role in the development of the increased alcohol tolerance associated with addiction[27]. The significance of this for the pathogenesis and treatment of AWS is not known.

Endocrine Changes: Abnormalities of endocrine function have been observed for several systems during the AWS. These are thyroid function, prolactin, FSH and LH, growth hormone, etc. It is difficult, because of inconsistencies to determine the validity of many of these observations, which may be secondary to other physical abnormalities consequent upon alcoholism. There are two abnormalities about which there is little doubt.

a) Hypothalamic-pituitary-adrenocortical activity is affected by alcohol. When certain concentrations are reached there is an increase in ACTH, and cortisol levels. Undoubtedly the effect is mediated by the hypothalamus and ACTH releasing hormone, whether this

is attributable to neural transients from higher CNS centres or part of a general stress reaction is not known. When alcohol is withdrawn and a significant AWS occurs these values remain high returning to normal in most alcoholics after three to four weeks[27]. Khan et al[28] have recently reviewed the literature on the results of the dexamethasone suppression test in alcoholism[28]. They found wide variation in the results, but of 170 alcoholics overall, 20% were non-suppressors in the first two weeks following withdrawal and 5% after more than two weeks abstinence.

b) Vasopressin: The old view that alcohol withdrawal was associated with a wet brain (hence the term 'drying out') was discounted when electrolyte studies became available and dehydration was believed to be the general rule. This view was challenged some 17 years ago[29], but it was not until recently when measurement of vasopressin became possible that these older views appeared correct. Alcohol reduces vasopressin secretion producing a diuresis and in alcohol withdrawal a rebound occurs with a rise in both vasopressin and its neurophysin precursors[30,31]. The presence of increased free water in the brain during alcohol withdrawal has been confirmed by NMR studies[32]. (This confirms the observation that patients with periorbital oedema have more severe withdrawal symptoms).

Neurotransmitters: It is difficult to generalize about the roles of the major neurotransmitters in AWS. The findings are varied both in terms of concentration and turnover. Many widely accepted findings such as reduced 5H-T levels do not invariably hold for all studies. These have been admirably summarized by Blum[33].

Benzodiazepines remain the drugs of choice in the management of AWS and the only new developments have been suggestions for the refinement of their use[34]. Drugs which have been recently used to manage AWS are Valproic acid[35,36] and carbamazepine[37]. More clinical trials of both are required, before their efficacy can be fully determined. However, both these drugs are of interest, Valproic acid because it is a GABA agonist and carbamezapine because it appears to reduce the dysthymia which is often a residual state after recovery from the AWS.

Other Treatments

More time could have been given to the neuropsychological, electrophysiological and CAT scan changes observed in alcoholism. From the point of view of biological treatments for the retention of abstinence once it has been procured there is little new to be said. New biological preventive treatments will come I believe from a better understanding of the physical process of addiction.

I would like to conclude with a quotation from Blum, "Alcohol is a very complex agent having a multitude of central nervous system actions... it is not simply an anaesthetic agent, nor is it a protein denaturant, nor is it a toxic pathogen; nor is it a false neurotransmitter; nor is it an

opiate receptor agonist; nor an antagonist; nor is it a releaser of biogenic amines and other neuropeptides...; nor is it an intoxicant; nor is it an addictive agent leading to devastation of the individual. It is not one of these, it is all of them."

REFERENCES

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (Third Edition). Author: Washington, D.C., 1980.

2. Mulford, H.A. and Fitzgerald, J.L. On the validity of the Research Diagnostic Criteria, the Feighner Criteria, and the DSM-III for diagnosing alcoholics. J Nerv Ment Dis, 1981, 169, 654-8.

3. Boyd, J.H., Weissman, M.M., Thompson, W.D. and Myers, J.K. Different definitions of alcoholism, I: impact of seven definitions on prevalence rates in a community survey. Am J Psychiatry, 1983, 140, 1309, 13.

4. Boyd, J.H., Derr, K. and Grossman, B. et al. Different definitions of alcoholism, II: a pilot study of 10 definitions in a treatment setting. Am J Psychiatry, 1983, 140, 1314-7.

5. Magruder-Habil, K., Jarris, K.E. and Fraker, G.G. Validation of the Veterans Alcoholism Screening Test. J Stud Alc, 1982, 43, 910-26.

6. Morgan, M.V. Markers for detecting alcoholism and monitoring for continued abuse. Pharacol Biochem Behav, 1984, 13 (Suppl 1), 1-8.

7. Chick, J., Krietman, N. and Plant, M. Mean cell volume and gamma-glutamyltranspeptidase as markers for drinking in working men. Lancet, 1981, 1:1249-55.

8. Beresford, T., Low, D., Hall, R.C., Adduci, R. and Goggans, F. A computerized biochemical profile for detection of alcoholism. Psychosomatics, 1982, 23, 713-4, 719-20.

9. Bernadt, M.W., Mumford, J. and Murray, R.M. A discriminant-function analysis of screening tests for excessive drinking and alcoholism. J Stud Alc, 1984, 45, 81-86.

10. Reich, T., Robins, L.N., Woodruff, R.A., Taibleson, M., Rich, C. and Cunningham, L. Computer-assisted deviation of a screening interview for alcoholism. Arch Gen Psychiatry, 1975, 32, 847-52.

11. Nasrallah, H.A. Laterality shift in alcoholic males. Biol Psychiatry, 1983, 18, 1065-7.

12. McShane, D. and Willenbring, M.L. Differences in cerebral asymmetries related to drinking history and ethnicity. J Ner Ment Dis, 1984, 172, 529-32.

13. Harburg, E. Handedness and drinking smoking types. Percept Motor Skills, 1981, 52, 279-82.

14. Mirabile, C.S., Glueck, B.C. and Hedberg, C.L. Motion sickness susceptability population profiles in a general psychiatric population and in heroin addicts and alcoholics. Int J Addict, 1981, 16, 1289-93.

15. Mellor, C.S. and White, H.P. Taste aversions to alcoholic beverages conditioned by motion sickness. Am J Psychiatry, 1978, 135, 125-6.

16. Steudig-Lindberg, G. and Rudy, N. Stepwise regression analysis of an intensive one-year study of delirium tremens. Acta psychiatr Scan, 1980, 62, 273-93.

17. Hemingsen, R. and Kramp, P. Effects of acute ethanol intoxication and ethanol withdrawal on magnesium and calcium metabolism in the rat. Psychopharmacology, 19809; 67, 255-259.

18. Davis, V.E. and Walsh, M.J. Alcohol amines and alkaloids. A possible basis for alcohol addiction. Science, 1970, 167, 1005-7.

19. Gallant, D.M. Alcoholism, naloxone and endorphin. Alcoholism Clin Exp Res, 1981, 5, 469-71.

20. Gold, M.S., Pottash, A.C., Ammito, W.J., Extein, I. and Kkleber, H.D. Loxefindine, a clonidine alalogue effective in opiate withdrawal. Lancet, 1981, 1:992-3.

21. Washton, A.M. and Resnick, R.B. Clonidine for opiate detoxication: Out-patient clinical trials. Am J Psychiatry, 1980, 137, 1121-2.

22. Walmider, J., Balldin, J., and Bokstrom, K. et al. Clonidine suppression of the alcohol withdrawal syndrome. Drug Alc Depend, 1981, 8, 345-8.

23. Hemmingsen, R., Clemmesen, L. and Barry, D.I. Blind study of the alpha-adrenergic agonists clonidine an lofexidine of alcohol withdrawal in the rat. J Stud Alc, 1984, 45, 310-15.

24. Eisenberg, S. Cerebral blood flow and metabolism in patients with delirium tremens. Clin Res, 1968, 16, 71.

25. Berglund, M. and Risberg, J. Regional cerebral blood flow during alcohol withdrawal related to consumption and clinical symptomatology. Act Neurol Scan (Suppl 64) 1977, 56, 480-1.

26. Gillman, M.A. and Lichtigfeld, F.J. Receptor hypothesis of the alcohol withdrawal state. Adv Biochem Psychopharmacol, 1983, 37, 405-15.

27. Stokes, P.E. Endocrine disturbances associated with alcohol and alcoholism. In: Pattison, E.M., Kaufman, E., Eds. Encyclopedia handbook of alcoholism. New York: Gardner Pre, 1982, 311-324.

28. Khan, A., Cirauto, D.A., Nelson, W.H., Becker, J.T., Nies, A. and Jaffe, J.H. Dexammethasone Suppression test in recently detoxified alcoholics: clinical implications. J Clin Psychopharmacol, 1984, 4, 94-7.

29. Beard, J.D. and Knott, D.H. Fluid and electrolyte balance during acute withdrawal in chronic alcoholic patients. JAMA 1968, 204, 135-9.

30. Eisenhoffer, G., Whiteside, E., Lambie, D., Johnson, R. Brain water during alcohol withdrawal. Lancet, 1981, 2, 923.

31. Legros, J.J., Deconinck, D. and Willems, B. Increase of neurophysin II serus levels in chronic alcoholic patients: relationship with alcohol consumption and alcoholism markers during alcohol withdrawal. J Clin Endocrinol Metab, 1983, 56, 871-5.

32. Besson, J.A., Glene, A.I., Foreman, E.I. and MacDonald, A. Nuclear magnetic resonance in alcoholic cerebral disorder and the role of vasopressin. Lancet, 1981, 2, 923-4.

33. Blum, K. Neurophysiological effects of alcohol. In: Pattison, E.M., Kaufman, E., eds. Encyclopedia handbook of alcoholisms. New York: Gardner Press, 1982, 105-34.

34. Sellers, E.M., Naranjo, C.A., Harrison, M., Devenyi, D.P., Roach, C., Sykora, K. Diazepam loading: simplified treatment of alcohol withdrawal. Clin Pharmacol Ther, 1983.

35. Caraffa, F. and Fion, I. L'impego del DPA (depaken) nelle complicanze psichishe e neurologiche dell'etillisimo. Neuropsichiatr, 1979, 65, 225-233.

36. Lambie, D.G., Johnson, R.H., Vijayasenan, M.E. and Whiteside, E.A. Sodium valproate in the treatment of the alcohol withdrawal syndrome. Aust NZ J Psychiatry, 1980, 14, 213-5.

37. Ballenger, J.C. and Post, R.M. Carbamazepine in alcohol withdrawal syndromes and schizophrenic psychoses. Psychopharmacol Bull, 1984, 20, 572-84.

FAMILIAL ALCOHOLISM AND THE IDENTIFICATION OF HIGH RISK SUBJECTS

MAURICE DONGIER, M.D.
Professor and Chairman
Department of Psychiatry
McGill University

ABSTRACT

In addition to the psychosocial high risk factors, biological predisposition to alcoholism has been studied in recent years: adoption studies in particular have pointed out that alcoholism runs in families not so much because of broken homes and disturbed childhoods, but because it is genetically transmitted to male first degree relatives. This understanding of familial alcoholism can lead to:

a) New strategies to screen for alcoholism in the families of known alcoholics, to attempt to break the denial barrier and to carry out early counselling in high risk subjects.

b) New researches to assess the predictive value of a number of biological explorations (neurochemical, neurophysiological, neuropsychological).

Familial Alcoholism: Basis for Identification of High Risk Subjects

The need for prevention of alcoholism has been emphasized repeatedly by agencies such as the World Health Organization, but there is no organized approach to recruiting potential alcoholics or alcoholics into treatment at an early stage. Most alcoholics enter treatment (if they do) after 15 or 20 years of abuse. This means usually that the individual has accumulated severe medical, psychological, social and economic problems. Medical treatment focuses at the present time on secondary prevention (e.g., disulfiram) and tertiary prevention (detection and treatment of complications).

Alcohol abuse is too often considered as an essentially socio-political issue, which can be dealt with by various methods such as increased taxes on alcoholic beverages, random roadside testing of blood alcohol content, increased legal age of drinking, banning of publicity on alcohol, and other such measures. However, only a small percentage of alcohol consumers are alcohol abusers: apart from general social measures, a more effective strategy may be to focus on the smaller number of potential alcohol abusers.

Background and Rationale of a Strategy

Any lay observer of human behaviour after ingestion of small doses of alcohol can note a wide variety of reactions: some individuals become happier, some more aggressive, some more introvert, some more extrovert. There are inter-individual differences as well as differences between ethnic groups. Pleasant or unpleasant mood changes, marked differences in the development of dependence and tolerance over time, all of these are based on variations in metabolism of alcohol, largely inherited. Knowledge of these factors can go a long way in the understanding of the etiology of alcoholism and in the descriptions of subtypes of drinking behaviour, no longer characterized exclusively by behavioural phenomenology and descriptive case histories, but also by underlying biological mechanisms of predisposition to alcohol abuse and dependence.

A number of studies carried out in Scandinavia and North America have shown beyond doubt that the risk of alcoholism in first degree relatives (parents, children, siblings) of alcoholic probands is about four times higher than in the general population. On the other hand, the risk is higher to male (25%) than to female (5%) relatives. Children of alcoholics are more likely to be alcoholic themselves than children of non-alcoholics, even when they have been adopted away early in life by non-alcoholic families (Goodwin et al., 1974; Cotton, 1979; Cadoret et al., 1980). These kinds of studies have helped distinguish between "nurture" and "nature" in the etiology of alcoholism.

Goodwin et al. studied a sample of 55 Danish males adopted in the first few weeks of life by non-relatives. Their biological parents were alcoholics. They were matched by age with a control group of 78 male adoptees whose biological parents were not alcoholics. The occurrence of alcoholism among the sons of alcoholics was about four times higher than in the sons of non-alcoholics, although there was no exposure to the alcoholic biological parent after the first few weeks of life. This was not accompanied by an increased risk of developing psychiatric disorders other than alcoholism. When compared with sons of alcoholics raised by alcoholic biological parents, exposure to the alcoholic parent did not appear to augment the increased genetic susceptibility.

Goodwin suggests that alcoholism be separated into two categories: familial and non-familial. "The former would be manifested by alcoholism in blood relatives, by early onset of alcoholism, and by a rapid development and fulminating course of the disease, requiring early treatment," the non-familial would have opposite characteristics. In most series, about one-half of patients do not report having alcoholic relatives.

Describing the weight of genetic factors does not mean neglecting the impact of psychological and social events but rather leads to more precise assessment of their relative importance: heritability is the proportion of variance of a characteristic that can be attributed to genetic factors. It is assessed in particular by twin studies, comparing monozygotic and dizygotic twins. If a variable were entirely under genetic influence, 100% of monozygotic twins, sharing the same genes, would be affected -

heritability would be 1.00. It proves to be complex and multifaceted when the many variables of drinking behaviour are considered. For instance, heritability of loss of control (inability to cease drinking) a major symptom of alcoholism, has been found in Finnish twin studies (Partanen et al., 1966) to be very much age dependent. Heritability of loss of control was negligible (0.07) for older twins but high (0.54) for younger twins.

Similarly, for total alcohol consumption (Kapprio et al., 1978) the highest heritability was among males for the age range 20-24 years, (0.53) but by age 70-74, heritability had declined to 0.04. Overall heritability for female consumption was rather lower 0.25.

In addition to adoption and twin studies, one usually turns in human genetics to the investigation of genetic markers, but this has been so far largely unproductive in the field of alcoholism. Four types of genetic markers have been used:

1. Blood groups and serum proteins;
2. Color vision defects;
3. Taste sensitivity of subjects for solutions of phenylthiourea; and
4. Secretion or non secretion of ABH blood group substance in the saliva.

There is no hard evidence to support the contention that any of these markers is associated with predisposition to alcoholism:

1. ABO blood groups, MNS or C_3 systems, rhesus system, have been investigated with contradictory findings.

2. The evidence regarding colour vision findings is confusing. Often colour blindness is of nutritional or other non-genetic origin.

3. Phenylthiocarbamide (PTC) is one example of phenylthiourea; this compound has been used in most studies of taste sensitivity in alcoholics. An excess of non-tasters has been found in alcoholics as in male schizophrenia and in adenomatous thyroid goiters. However, it is not clear whether the loss of taste sensitivity is a result of alcohol intake or is caused by genetic mechanisms.

4. Secretion of the antigens A, B and H of the ABO system is controlled by a single pair of allelomorphic genes. Thirty-three percent of alcoholics have been found non-secretors, for group A, but the explanation that alcohol intake influences secretion has not been excluded.

What Is Inherited?

Obviously, what is inherited is a vulnerability to a number of environmental factors, (a major one being the availability of alcohol). Research into finding traits or markers which can identify early in life vulnerable individuals, i.e., potential compulsive drinkers, is a recently developed area in alcohology.

At least in an important proportion of cases, what is transmitted genetically involves the mechanisms of tolerance or at least lack of tolerance. More than half of the Orientals, in response to a small dose of alcohol, develop a flushing response, a feeling of malaise and a disinclination to continue drinking, as if they had been administered disulfiram (Antabuse). This intolerance is genetic in nature and related to the enzymes which control the metabolism of alcohol.

Two main categories of enzymes have been studied: alcohol dehydrogenases (ADH) and acetaldehyde dehydrogenases (ALDH).

Numerous varieties (isozymes) of ADH have been described (Vallée and Bazzone, 1983) in human liver and in a number of animal species. They are directly related to genetic coding, and in different proportion within ethnic groups. An obvious hypothesis comes to mind: some variants of ADH or ALDH could be identified in certain families or ethnic groups and correlated with vulnerability to alcoholism. Unfortunately, their measurement in peripheral blood raises problems, and most of the work published so far has been carried out post-mortem on liver specimens. Atypical liver alcohol dehydrogenase (ADH) is very frequent among the Japanese (Stamatoyannopoulos et al., 1975). Analysis of 40 autopsied liver specimens showed that 34 had an atypical ADH phenotype. Also atypical ADH were found by von Wartburg (1968) in 20% of Swiss and 4% of English livers. The fact that possibly 80% of Japanese carry an atypical liver ADH suggests that the low rate of alcoholism in Orientals may be linked with increased acetaldehyde formation in individuals carrying an atypical ADH gene.

ADH on the one hand, ALDH on the other hand, regulate blood acetaldehyde level. High levels of acetaldehyde may be responsible for unpleasant reactions to drinking, such as flushing, nausea, hypotension, etc. Low prevalence of alcoholism in Orientals has been attributed to a high plasma acetaldehyde. On the other hand, in Caucasians, high plasma acetaldehyde has been found in young men with alcoholic relatives (Schuckit and Rayses, 1979), two rather contradictory findings.

It is very likely, though not so clearly established, that the contrary, namely, excessive tolerance leading to increased consumption in order to reach a "high", is also genetically determined in a number of subjects. This is the human equivalent of the observations in animals of increased tolerance and predisposition to abuse obtained by selective breeding.

How alcohol produces euphoria is not known precisely. Does alcohol affect monoamines or other brain chemicals to change mood and cognition? Data in this area are not convincing. A hypothesis has been offered on the basis of the combination of acetaldehyde with dopamine generating tetrahydropapaveroline, an endogenous precursor of morphine. This endogenous synthesis of opiate-like compounds might underlie the addictive properties of alcohol in susceptible individuals (Blum et al., 1982).

If alcoholics have higher levels of acetaldehyde than non-alcoholics after ingestion of alcohol, possibly they maintain their drinking in order to overcome the unpleasant effects of acetaldehyde. This model of addiction has been suggested by Wikler (1973): euphoric effects are followed by dysphoric effects, leading to repeated attempts to overcome the latter by increasingly heavy use of the drug.

Assessment of plasma acetaldehyde is still a very controversial issue, and practically all studies carried out so far do not have a high level of credibility. The high risk group of children of alcoholics is certainly worthwhile studying closely.

A Possible Strategy

The families of alcoholics should be cognizant of the familial nature of many cases of alcoholism. It is possible to offer an assessment of drinking habits and predisposition to alcohol abuse to blood relatives of identified alcoholics. Children of alcoholics, as soon as they reach drinking age, can be taught to recognize early symptoms, and be made aware that their chance of becoming alcoholic is four times greater than for children of non-alcoholics. Although no reliable biological marker of predisposition is available so far, they can be offered to volunteer for a number of tests for research purposes, and be given feedback on the findings. Whatever the results of these, they should exercise special caution in their use of alcohol. If they are found by systematic screening as already on their way to develop the addiction, early intervention can take place and they may be oriented to the treatment of choice in their case. At the same time, they should be screened for psychiatric symptoms, to check the important finding by Vaillant (1983) that depression, anxiety and antisocial behaviour are symptoms caused by alcoholism more often than the reverse.

The possibility has been raised by Goodwin that disulfiram (antabuse) may be a more useful treatment for familial alcoholics than for other alcoholics.

Intensive study of the biochemical and the neurophysiological characteristics of these high risk subjects may be an important way of discovering a treatment of alcoholism which would no longer be purely psychosocial but biochemical and based on an understanding of the mechanisms responsible for the addiction. (Tests carried out on an experimental basis, have included so far plasma acetaldehyde, endogenous opiates, EEG under alcohol challenge, psychological testing to assess anhedonia, etc.)

A one-year pilot study has shown that it is feasible to reach this target population. This can be achieved in particular through contacts with self-help groups such as Al-Anon (spouses of alcoholics) and Al-Ateen (adolescent children of alcoholics). Contacts can also be established within a systematic fashion with blood relatives of alcoholics through rehabilitation centers, at the time of residential treatment of the probands, or during follow-up treatment.

Conclusion

Popularization of the genetic factors in the etiology of alcoholism, and the family approach for the identification of high risk subjects, may raise resistances, in particular, a conscious or unconscious non-sequitur: if alcoholism is hereditary, it cannot be treated. Experience shows that it is fairly easy to overcome such illogical reactions: most predisposed individuals can accept that heredity is predisposition, not implacable fate, and that many hereditary illnesses are treatable. Moreover, if as in the McGill-Douglas Hospital Alcohol Research project, high risk subjects are offered a number of investigations (psychological, neurophysiological, biochemical), to help in the search for mechanisms of predisposition, their motivation to take some steps to prevent the development of alcoholism may be enhanced.

REFERENCES

Blum, K., DeLallo, L., Briggs, A.H. and Hamilton, M.G. (1982): Opidoid responses of isoquinoline alkaloids (TIQS). In: Beta-Carbolines and Tetrahydroisoquinolines, Eds. F. Bloom, G. Barchas, M. Sandler and E. Usdin, New York, Alan R. Liss, pp. 387-398.

Cadoret, R.J., Cain, C.A. and Grove, W.M. (1980): Development of alcoholism in adoptees raised apart from alcoholic biological relatives. Archives of General Psychiatry, 37, 561-563.

Cotton, N.S. (1979): The familial incidence of alcoholism: A review. Journal of Studies in Alcohol, 40 (1), 89-116.

Goodwin, D.W., Schulsinger, F., Morr, N., Hermansen, L., Winokur, G. and Guze, S.B. (1974): Drinking problems in adopted and non adopted sons of alcoholics. Archives of General Psychiatry, 31, 164-169.

Kappiro, J., Sarna, S., Koskenvuo, M. et al. (1978): The Finnish Twin Registry, Baseline Characteristics Section II, Helsinki, University of Helsinki Press.

Partanen, J., Bruun, K. and Markkanen, T. (1966): Inheritance of drinking behaviour - Rutgers University Centre of Alcohol Studies, New Brunswick, New Jersey.

Schuckit, M.A. and Rayses, V. (1979): Ethanol ingestion: differences in blood acetaldehyde concentrations in relatives of alcoholics and controls. Science, 203, 54-55.

Stamatoyannopoulos, G., Chen, S.H. and Fukui, M. (1975): Liver alcohol dehydrogenase in Japanese: high population frequency of atypical form and its possible use in alcohol sensitivity. American Journal of Human Genetics, 27, 789-796.

Vaillant, G. (1983): The natural history of alcoholism, Harvard University Press, Cambridge, Mass.

Vallée, B.L. and Bazzone, T.J. (1983): Isozymes of human liver alcohol dehydrogenase. In: Isozymes: Current Topics in Biological and Medical Research, (Scandelios, J.G., Rattazzi, M.C. and Whitt, G.S., Eds.), Vol. 8: Cellular Localization Metabolism and Physiology, pp. 219-244. New York: Liss.

Von Wartburg, J.P. and Schurch, P.M. (1968): Atypical human liver alcohol dehydrogenase. Annals of New York Academy of Sciences, 151, 936-947.

Wikler, A. (1973): Dynamics of Drug Dependence: Implications of a Conditioning Theory for Research and Treatment. Archives of General Psychiatry, 28, 611-616.

A CRITICAL REVIEW OF DIAGNOSTIC AND THERAPEUTIC DEVELOPMENTS: DIAGNOSIS AND PSYCHOLOGICAL THERAPY

R.C. BOWEN, M.D.
Department of Psychiatry
University of Saskatchewan

Nosology

Agoraphobia is puzzling in many ways. In clinical psychiatric practice, it is the most common and distressing of the anxiety disorders. "At the simplest level, people disagree about how to pronounce it (whether to emphasize the first syllable or the second; Webster's Dictionary prefers the first)." (Thorpe, Burns, Smith & Blier, 1984) The Diagnostic and Statistical Manual of the American Psychiatric Association (DSM-III) lists 4 phobic disorders including agoraphobia with panic attacks and agoraphobia without panic attacks. Panic disorder is listed under "Anxiety States" along with Generalized Anxiety Disorder and Obsessive-Compulsive Disorder (American Psychiatric Association, 1980). This does not exhaust the categories in DSM-III that pertain to anxiety. Sheehan and Sheehan (1983) have pointed out that there are over 20 categories in DSM-III that are anxiety related. Combinations of some categories are allowed and other combinations are not allowed. We will consider only two of the conditions listed in the chapter titled Anxiety Disorders in DSM-III, panic disorder and agoraphobia with panic attacks.

One of the proposed revisions to DSM-III includes dropping the term "agoraphobia" (Spitzer & Williams, 1984). This is a progressive step because of the many misconceptions associated with the term, not the least of which was the definition. The literal definition of agoraphobia has been the "fear of open places" and the stereotyped description was that of the "housebound housewife." Both of these are incorrect. In the proposed revision, prominence has been given to "panic disorder," and the phobias have then been categorized according to the degree of restriction of activities. Panic disorder with extensive phobic avoidance would cover the old category of agoraphobia with panic attacks. It is recognized that symptoms very much resembling panic can occur in most phobic patients on exposure to the phobic stimulus. The proposed revision clarifies that spontaneous panic attacks must occur at times when they are unexpected, i.e., many patients report panic attacks when they are completely relaxed at home with no apparent relevant stimulus, and some patients report awakening while experiencing a panic attack. The symptoms and worst fears of patients with agoraphobia and panic attacks have been listed by various authors including Thorpe and Burns (1983). These fears are not of a particular location but of the consequences of fear itself.

A second proposed revision has been the removal of the heirarchical distinction between major depression and panic disorder. Under DSM-III, panic disorder would only be diagnosed if the panic attacks were separate from major depression. This partly arises from the idea that the patient

should only have one disease. Roth (1978) has illustrated implicit hierarchies in psychiatric classification systems. Depression, and in fact most other syndromes, have preceded anxiety in hierarchies, and I suspect also in research priorities. Spitzer and Williams (1984) have commented that many issues that should be seen as differential diagnostic issues, have been seen as heirarchical issues. The topic of the distinction between depression and anxiety cannot be covered here, but is obviously very relevant to biological researchers (Leckman, Weissman, Merikangas, Pauls and Prusoff, 1983)(Kerr, Roth, Schapira & Gurney, 1972).

A Brief Review of the Epidemiology

The Epidemiologic Catchment Area Studies have determined the lifetime prevalence rates for the common DSM-III disorders (Robins et al., 1984). The anxiety disorders are very common, of the same order of prevalence as the affective disorders and alcoholism, about 10% each. It is also apparent that the phobic disorders (including both agoraphobia and simple phobia), were diagnosed much more frequently in Baltimore compared with New Haven and St. Louis. The six month prevalence rates for specific anxiety disorders from the New Haven site have been published. Even if simple phobias are disregarded, panic and agoraphobia have a combined prevalence of 3.8/100, so these are common disorders (Leaf, Weissman, Myers, Tischler & Holzer, 1984).

Results from these studies have confirmed that anxiety disorders are more common among women, and are most common in the 25-44 age group. Several studies have shown a bimodal age of onset, but Sheehan, Sheehan & Minichiello (1981) found a peak in the 20's which appeared unimodal. They suggested that the bimodal peak found in other studies was an artefact of small numbers. Their data did demonstrate that panic attacks can start in childhood. Several studies of inpatient alcoholics, including one done by us in Saskatoon, have found that one-quarter of inpatient alcoholics suffer from severe phobias. In general, the anxiety disorders tend to develop several years before excessive drinking. We suggested that the whole topic of anxiety disorders and alcoholism deserves more research attention (Bowen, Cipywnyk, D'Arcy & Keegan, 1984).

Stability of Prevalence and Prognosis

Anxiety as well as depression tends to be more common among women. The stability of these findings over 18 years, from 1952 to 1970, has been investigated in the Stirling County Study (Murphy, Sobol, Neff, Oliver & Leighton, 1984). In 1970, rates of these disorders were still higher in women, but men and women were more similar, and the peak in older women seems to have shifted to younger women. The overall rates have tended to remain stable over 18 years despite marked changes in social conditions. The long term prognosis of these disorders has been investigated in several studies from Iowa. A cohort of more than 100 patients with panic disorder was traced over 40 years later. There was a control group of patients with unipolar affective disorders. It was found that the observed deaths

exceeded the expected deaths in the panic disorder patients, particularly in the fourth decade of the study, and that this difference was more striking in men. The excess was accounted for mainly by death by unnatural causes, predominantly suicide, and circulatory system disease (Coryell, 1984).

The natural course of these disorders has been a controversial issue over the years, but the claim of a high spontaneous remission rate has been supported by studies from Iowa (Rachman & Wilson, 1980). In one study, a group of 80 patients, who had been seen on a consultation service but had not been offered formal psychiatric therapy, were followed for five years. The symptoms developed at the mean age of 29 years and the patients were seen at a mean age of 36 years so they had had symptoms an average of seven years. The results confirmed that approximately two-thirds of the patients improved substantially. It is not clear why they had not improved before being seen (Noyes & Clancy, 1976). This study has been repeated on a larger group of patients compared with surgical control patients and followed after six years (Noyes, Clancy, Hoenk & Slymen, 1980).

It is apparent that there are substantial numbers of chronically disabled patients who present to specialized clinics. Advocates of the spontaneous remission theory would claim that the presence of chronic cases does not refute the argument that most patients improve within 2 years. The issue is further confounded by the follow-up studies from the Newcastle group who followed an older group of patients in their 40's and found that anxiety symptoms tended to be associated with a poorer prognosis. They mention that treatment for anxiety was not as readily available in the 1960's as treatment for depression, and that this might have contributed to the poorer prognosis (Kerr et al., 1972).

Most of the prognostic studies have found that a worse prognosis is associated with childhood "neurotic" conditions (Kerr et al., 1972). This raises the issue of prevention, since several authors have mentioned that even though clear panic attacks often start in the early 20's, many patients report childhood anxiety symptoms (Sheehan & Sheehan, 1983). Of relevance to this is the suggestion by Gittleman (1984) that school avoidance in adolescents is associated with the later development of agoraphobia.

Genetics

Family history studies from the University of Iowa have confirmed the increased risk of anxiety disorders among relatives of patients with panic disorder. The risk among first degree relatives is about 18% and among second degree relatives about 9%. There seems to be an increased risk for alcoholism among first degree relatives, but the findings on depression in relatives are inconclusive (Noyes, Clancy, Crowe, Hoenk & Slymen). Twin data have confirmed that panic disorder is five times as frequent in monozygotic co-twins as in dizygotic co-twins. Both a dominant gene model with variable penetrance and a polygenic model seem to fit the data (Torgerson, 1983).

Models of Development of the Phobic Syndrome

Klein (1981) has proposed the best known psychiatric theory on the development of agoraphobia. The first spontaneous panic attack is followed by recurring episodes. As time progresses with an increasing number of panic attacks, general discomfort and anticipatory anxiety increase. Help-appeal behaviour, i.e., visiting emergency rooms decreases. Avoidant and dependent behaviour, in other words phobias, increase. This is the model used in most treatment programs.

Sheehan (1982) has proposed a different model which focuses on diagnostic issues. He describes seven stages in the development of the syndrome culminating in eventual depression. He questions the tendency based on cross-sectional diagnoses to identify seven different disorders as the patient progresses through the phases. One certainly sees patients who progress through phases of this scheme but certainly not everyone progresses through all the phases, although Sheehan has not implied this in his model.

Mathews, Gelder & Johnston (1981) in England, have described a more comprehensive model. This model takes into account factors in early family development, e.g., maternal over-protection has been investigated (Parker, 1983). Increased non-specific stresses are a part of the model and the majority of patients mention some stress before the onset of panic attacks. Whether this is a valid observation or hindsight is not yet clear. This leads to elevated levels of general anxiety culminating in acute panic attacks if a person is genetically predisposed. Clearly, however, not everyone who develops panic attacks develops phobic avoidance behaviour and Mathews et al. mention an external locus of control and attribution of the attacks to external stimuli as factors in the development of avoidance.

Clearly, each of these models which are not mutually contradictory, make some valuable contributions and there are numerous hypotheses for future research embedded in each of the models.

Biological Theories of Panic Attacks

Four topics of recent interest will be mentioned.

In 1976, Wooley suggested that there was an association between the mitral valve prolapse syndrome (MVPS) and functional disorders with cardiac complaints, such as panic disorder. Since then, at least six studies have reported that about one-third of patients with agoraphobia and panic attacks suffer from the MVPS. This sparked interest in the MVPS as a biological marker for panic disorder. We have done a carefully controlled study with the Cardiology Department in Saskatoon and failed to find an association. This finding is consistent with the more carefully controlled studies. The last chapter has not been written on this purported association, but it does appear that sampling biases could have accounted for the results in the positive studies (Bowen, Orchard, Keegan & D'Arcy, 1985).

The monoamine oxidase (MAO) enzyme is of particular interest because of the therapeutic efficacy of MAO inhibitors in these patients. In a preliminary study, we have found increased MAO levels in a group of agoraphobics compared with depressed and normal controls. The level of para-hydroxyphenylacetic acid was lower in the agoraphobics and depressed controls compared with normal controls. This finding is consistent with the results of several other studies. It is not clear, at present, what the meaning of this finding is in terms of the etiology of the disorder (Yu, Bowen, Davis & Boulton, 1983).

Brief mention should be made of the benzodiazepine receptor. These receptors are found throughout the CNS but particularly in cortical areas. They are linked with GABA receptors which are closely associated with chloride channels. The benzodiazepines have marked anxiety reducing properties, and bind with high affinity to these receptors. During the past three years, antagonists of the benzodiazepine effect have been identified (RO 15-1788), and active antagonists (the beta-carbolines) which produce panic type symptoms in animals and man have been studied (Skolnick & Paul, 1983).

Finally, there is increasing evidence linking noradreneric (NA) function to fear and anxiety. Detailed studies have been done on the locus ceruleus, a small brain stem nucleus which supplies most of the NA connections to the limbic system. Electrical stimulation, benzodiazepine withdrawal and caffeine, increase locus ceruleus firing. Many drugs that reduce anxiety in humans, reduce firing of the locus ceruleus. It is not currently known how the benzodiazepine receptor complex and the NA system interact (Uhde et al., 1984).

Treatment

In 1981, the Australian and New Zealand Quality Assurance Project reviewed treatment outcome studies on agoraphobia. One hundred and forty-five articles on treatment were reviewed and 25 studies were identified which presented results in a form that allowed the calculation of an effect size. Overall, the mean effect size was 1.5, indicating that the average patient after treatment was better off than some 90% of untreated patients. There is a large literature on the psychological treatment of agoraphobia, and most aspects of behaviour therapy for these conditions have been investigated.

One topic from the psychological treatment literature might be of particular interest to biological researchers. This is the question of whether psychological interventions affect the frequency and/or severity of spontaneous panic attacks. The high placebo response rates in the various drug studies, and the clear improvement demonstrated in the behavioural studies, suggest that behavioural treatments influence the frequency and severity of panic attacks. It is true that in most behavioural studies, the main target has been phobic avoidance (Marks et al., 1983). If panic

attacks are reduced in frequency and/or severity, this would provide an interesting model for the effects of psychological therapies on a process that is apparently biologically driven (Waddell, Barlow & O'Brien, 1984).

Medical Utilization

Finally, a few comments on medical utilization. The ECA studies have confirmed that patients with anxiety disorders are high utilizers of general and mental health services (Shapiro et al., 1984). In a study in Saskatchewan, we have divided a group of 100 agoraphobics into high utilizers and low utilizers. The high utilizers use up to 80 services per year on an average before treatment. All high utilizers who are referred decrease their utilization, but those who comply with a behavioural program show a greater fall in utilization than those who do not comply. Most of this treatment is carried out in groups by nurse therapists (Bowen, D'Arcy, South & Hawkes, 1984). Ginsberg, Marks and Waters (1984) have carried out a controlled study with similar results.

REFERENCES

American Psychiatric Association (1980). Diagnostic and statistical manual of mental disorders (3rd ed.), pp. 225-239. Washington, D.C.: Author.

Bowen, R.C., Cipywnyk, D., D'Arcy, C. and Keegan, D.L. (1984). Alcoholism, Anxiety Disorders and Agoraphobia. Alcoholism: Clinical and Experimental Research, 8, 48-50.

Bowen, R.C., D'Arcy, C., South, M. and Hawkes, J. (1984, October). The Effects of a Nurse Therapist Directed Behavioural Agoraphobia Treatment Program on Medical Utilization. Paper presented at the Canadian Psychiatric Association Meeting, Banff, Alberta.

Bowen, R.C., Orchard, R.C., Keegan, D.L. and D'Arcy, C. (1985). Mitral Valve Prolapse and Psychiatric Disorders. Manuscript submitted for publication.

Coryell, W. (1984). Mortality After Thirty to Forty Years: Panic Disorder Compared with Other Psychiatric Illnesses. In: L. Grinspoon (Ed.), Psychiatry Update: The American Psychiatric Association Annual Review, Vol. III, p. 463. Washington, D.C.: American Psychiatric Press.

Ginsberg, G., Marks, I., and Waters, H. (1984). Cost-benefit analysis of a controlled trial of nurse therapy for neuroses in primary care. Psychological Medicine, 14, 683-690.

Gittleman, R. (1984). Anxiety Disorders in Children. In: L. Grinspoon (Ed.), Psychiatry Update: The American Psychiatric Association Annual Review, Vol. III, pp. 410-418. Washington, D.C.: American Psychiatric Press.

Kerr, T.A., Roth, M., Shapira, K. and Gurney, C. (1972). The Assessment and Prediction of Outcome in Affective Disorders. British Journal of Psychiatry, 121, 167-174.

Klein, D.F. (1981). Anxiety Reconceptualized. In D.F. Klein, J. Rabkin (Eds.), Anxiety: New Research and Changing Concepts, p. 237, New York: Raven.

Leaf, P.J., Weissman, M.M., Myers, J.K., Tischler, G.L. and Holzer, C.E. (1984). Social factors related to psychiatric disorder: the Yale Epidemiologic Catchment Area Study. Social Psychiatry, 19, 53-61.

Leckman, J.F., Weissman, M.M., Merikangas, K.R., Pauls, D.L. and Prusoff, B.A. (1983). Panic Disorder and Major Depression. Archives of General Psychiatry, 40, 1055-1060.

Marks, I.M., Gray, S., Cohen, D., Hill, R., Mawson, D., Ramm, E. and Stern, R.S. (1983). Imipramine and Brief Therapist-Aided Exposure in Agoraphobics Having Self-exposure Homework. Archives of General Psychiatry, 40, 153-162.

Mathews, A.M., Gelder, M.G. and Johnston, D.W. (1981). Agoraphobia: Nature and Treatment, p. 47. New York: Guilford.

Murphy, J.M., Sobol, A.M., Neff, R.K., Oliver, D.C. and Leighton, A.H. (1984). Stability of Prevalence: Depression and Anxiety Disorders. Archives of General Psychiatry, 41, 990-997.

Noyes, R. and Clancy, J. (1976). Anxiety Neurosis: A 5-year Follow-up. The Journal of Nervous and Mental Disease, 162, 200-205.

Noyes, R., Clancy, J., Crowe, R., Hoenk, P.R. and Slymen, D.J. (1978). The Familial Prevalence of Anxiety Neurosis. Archives of General Psychiatry, 35, 1057-1059.

Noyes, R., Clancy, J., Hoenk, P.R. and Slymen, D.J. (1980). The Prognosis of Anxiety Neurosis. Archives of General Psychiatry, 37, 173-178.

Parker, G. (1983). Parental Over protection: A Risk Factor in Psychosocial Development, pp. 185-194. New York: Grune and Stratton.

Rachman, S.J. and Wilson, G.T. (1980). The Effects of Psychological Therapy (2nd ed.), pp. 24-49. Oxford: Pergamon.

Robins, L.N., Helzer, J.E., Weissman, M.M., Orvaschel, H., Gruenberg, E., Burke, J.D. and Regier, D.A. (1984). Lifetime Prevalence of Specific Psychiatric Disorders in Three Sites. Archives of General Psychiatry, 41, 949-958.

Roth, M. (1978). Psychiatric Diagnosis in Clinical and Scientific Settings. In: H.S. Akiskal, W.L. Webb (Eds.), Psychiatric Diagnosis: Exploration of Biological Predictors, p. 18. New York: Spectrum.

Shapiro, S., Skinner, E.A., Kessler, L.G., Von Korff, M., German, P.S., Tischler, G.L., Leaf, P.J., Benham, L., Cottler, L. and Regier, D.A. (1984). Utilization of Health and Mental Health Services: Three Epidemiologic Catchment Area Sites. Archives of General Psychiatry, 41, 971-978.

Sheehan, D.V. (1982). Current Perspectives in the Treatment of Panic and Phobic Disorders. Drug Therapy, 12, 49-63.

Sheehan, D.V. and Sheehan, K.H. (1983). The Classification of Phobic Disorders. International Journal of Psychiatry in Medicine, 12, 243-266.

Sheehan, D.V., Sheehan, K.E. and Minichiello, W.E. (1981). Age of Onset of Phobic Disorders: A Re-evaluation. Comprehensive Psychiatry, 22, 544-553.

Skolnick, P. and Paul, S.M. (1983). New Concepts in the Neurobiology of Anxiety. Journal of Clinical Psychiatry, 44, 12-19.

Spitzer, R.L. and Williams, J.B. (1984). Diagnostic Issues in the DSM-III Classification of the Anxiety Disorders. In: L. Grinspoon (Ed.), Psychiatry Update: The American Psychiatric Association Annual Review, Vol. III, p. 394. Washington, D.C.: American Psychiatric Press.

The Quality Assurance Project (1982). A Treatment Outline for Agoraphobia. Australian and New Zealand Journal of Psychiatry, 16, 25-33.

Thorpe, G.L. and Burns, L.E. (1983). The Agoraphobic Syndrome: Behavioural Approaches to Evaluation and Treatment, p. 25. Chichester: John Wiley and Sons.

Thorpe, G.L., Burns, L.E., Smith, P.J. and Blier, M.J. (1984). Agoraphobia: Research Developments and Clinical Implications. In: C.M. Franks (Ed.) New Developments in Behaviour Therapy: From Research to Clinical Application, p. 281. New York: Howorth.

Torgersen, S. (1983). Genetic Factors in Anxiety Disorders. Archives of General Psychiatry, 40, 1085-1089.

Uhde, T.W., Boulenger, J.P., Post, R.M., Siever, L.J., Vittone, B.J., Jimerson, D.C. and Roy-Byrne, P.P. (1984). Fear and Anxiety: Relationship to Noradrenergic Function. Psychopathology, 17 (Suppl. 3), 8-23.

Waddell, M.T., Barlow, D.H. and O'Brien, G.T. (1984). A Preliminary Investigation of Cognitive and Relaxation Anxiety of Panic Disorder: Effects of Intense Anxiety vs. 'Background' Anxiety. Behaviour Research and Therapy, 22, 393-402.

Wooley, C.F. (1976). Where are the diseases of yesteryear? Da Costa's syndrome, soldiers heart, the effort syndrome, Neurocirculatory asthenia--and the mitral valve prolapse syndrome. Circulation, 53, 749-751.

Yu, P.H., Bowen, R.C., Davis, B.A. and Boulton, A.A. (1983). Platelet monoamine oxidase activity and trace acid levels in plasma of agoraphobic patients. Acta Psychiatrica Scandinavica, 67, 188-194.

REFLECTIONS

Comments by ALASTAIR MUNRO, M.D.
who chaired one of the sessions in which
the preceding papers were presented

This morning's session is very heartening but has to be set against a disturbing reality.

I am militantly optimistic about the future of psychiatry as a therapeutic specialty, proud of many of its achievements, and yet frequently appalled by the apathy, and total lack of scientific curiosity or judgement of many of my psychiatric colleagues.

There are many internists and surgeons who never do research, but anyone in these specialties who wants to work in the advanced clinical centres is very likely to have to do research as a trainee to gain entry, and continue to do it to maintain credibility.

How different it is in psychiatry. Few of our trainees become familiar with research methodology, let alone research, but once they become specialists they encounter an anti-research bias, largely based on remuneration but also on prestige. As long as the beau-ideal in psychiatry is the therapist in solitary practice, tied to a method of therapy which limits him to a very circumscribed patient population, our attempts to promote the use and application of research are doomed.

The themes and presentations of this conference are excellent in intent and represent great hope for the future. Unfortunately they will sail clean over the heads of the great majority of psychiatric practitioners, and this is a great cause for concern. The practitioners who are attuned to what the participants of this conference are saying are in a small minority. We do not see a constant and productive two-way interchange between psychiatric research and psychiatric practice because so much of the latter is divorced from scientific thought.

I think we may see a great split in psychiatry before too long. If the forces of action are to gain the upper hand, we need a great deal of help, politically and medico-politically, to convert psychiatric practice to here-and-now terms. Of course psychotherapy and other psychological treatments will continue to be in our armamentarium but they have to be ancillary to the kinds of advances this conference is addressing.

If we cannot wrench our specialty into the present day, we will fail to recruit young people who are capable of productive research. Then we are like beggars waiting for crumbs to fall from other people's tables. How many of the marvellous new developments pertinent to psychiatry in the past 30 years really came from within our specialty?

So, I hope we can take what we are learning at this conference and bring it to our colleagues so that it may be usefully applied.

PART II. NEW BIOTECHNOLOGICAL DEVELOPMENTS

THE IMPACT OF MOLECULAR BIOLOGY

MOLECULAR BIOLOGY APPROACH TO STUDY OF THE BRAIN FUNCTION

M. CHRÉTIEN, M.D.
C. LAZURE, Ph.D.
N.G. SEIDAH, Ph.D.
Clinical Research Institute of Montreal

ABSTRACT

I have been asked to try to present to you a bird's-eye view of the impact of the molecular biology on the diagnosis and treatment of psychiatric disease.

The task is monumental and may be premature. I will try to demonstrate to you that there is hope that the new biotechnological developments might help at least to understand certain mechanisms of the brain functions. This will not necessarily bring new diagnostic tools and new developed drugs the next morning, the next week, the next month, or even the next year.

It will, however, unravel certain structures of basic components of the brain cells: the membranes and their lipid protein interactions, the axonal transports, the enzymatic systems of neuropeptides maturation, the genetic expression of brain proteins, the recognition signals for the growth of neural processes, the memorisation of all information by the brain, and its availability at any time ("at least when you don't suffer of memory loss"!!!)

The knowledge of the chemistry of the basic elements of the brain shall help to understand better the numerous interrelationships of these molecules in the trillions of connections contained in mammalian brain.

The molecular biology approach to the study of brain function was really initiated in the 1960's with the characterization of the biosynthetic pathways of the classical neurotransmitters discovered in the 1940's and the 1950's. We were then talking about a little more than 10 neurotransmitter systems which are, however, present everywhere in the brain. The late 1960's started the golden times of neuropeptides with the description of the chemical structures of TRF, LH-RF and somatostatin. Soon after came substance P, neurotensin, VIP, CCK, etc... The excitement came to its maximum with the discovery of the enkephalins and the endorphins. We were fortunate to be an active participant to this era which is still going on with tremendous speed. Our involvement has been mainly in the description of their biosynthetic pathways and the chemical characterization of human β-endorphin.

General Introduction

During the past few years, much attention has been focused on the study of the molecular interactions involved in the biosynthesis and segregation of secretory proteins. The development of techniques to isolate specific mRNAs and the availability of cell-free protein synthesis systems have stimulated the study of the primary translation products and their processing. We will present in the next pages the molecular interactions involved in the secretory process, and will summarize the present knowledge about the maturation of human pro-opiomelanocortin into β-endorphin, ACTH, MSH's (Chrétien et al., 1979a; Chrétien et al., 1979b) and a new N-terminal fragment which is a major secretory product (Seidah, Chrétien, 1981a; Seidah et al., 1981d). We will then apply this type of maturation to other models of pro-hormones and pro-peptides in order to demonstrate the secretory products which are or shall be expected. The conclusion will be that this type of maturation was foreseen in 1967 when we presented the hypothesis of pro-hormone using β-lipotropin as model (Chrétien, Li 1967).

Biosynthesis and Segregation of the Pro-Hormones

The biosynthesis of a secretory protein, as for any protein of the cell, starts in the cytosol by the formation of a complex between the ribosome, and mRNA, the initiator tRNA and a group of protein factors (Weissbach, 1979). The synthesis of the protein proceeds until the chain is about 60 or 70 amino acids long. Then the signal peptide that has emerged from the big subunit of the ribosome promotes the formation of a junction between the 60S subunit of the ribosome and the membrane of the endoplasmic reticulum. The exact mechanism by which the signal peptide initiates the binding of the ribosome to the membrane and the translocation of the nascent polypeptide chain through the membrane is not known. However, Blobel and Dobberstein (1975) first proposed that the interaction of the signal peptide with the membranes recruits some membrane protein receptors and causes their association to form a tunnel across the membrane. The tunnel is further stabilized by the binding of the big ribosomal subunit to the membrane. The newly formed peptide chain is then inserted in the tunnel and driven across the membrane by the sole force of the translational machinery. Once the translocation of the nascent chain is initiated, the signal peptide is cleaved by an endopeptidase called signal peptidase.

The examination of the primary structure of the signal peptides of different pre-proproteins listed in Figure 1 reveals some common characteristics: the signal peptides are 15 to 30 amino acids in length and are characterized by a region rich in hydrophobic amino acids (boxed in Figure 1) followed by a region containing small neutral amino acids. The role of the signal peptide is clear: to promote the ribosome binding to the endoplasmic reticulum membrane and to initiate the nascent polypeptide chain translocation across the membrane via the SRP receptor and the complex SRP-signal peptide.

The co-translational cleavage of the signal peptide is done by an endopeptidase located at the luminal side of the endoplasmic reticulum membrane. An enzyme from the bacterium Escherichia coli has been isolated (Zwizinski, Wickner 1980) and found to be a protein with a molecular weight of 39,000 daltons. The bacterial enzyme does not need a soluble factor to be active. The characterization of the eukaryotic enzyme is not yet done; it has been solubilized in detergent and its activity and specificity conserved (Jackson, Blobel 1980). The specificity of the enzyme is not very high. Examination of the primary structure of the pre-proteins around the cleavage site indicates no common sequence of amino acids. However, neutral or hydrophobic amino acids with small side chains are always found at the carboxyl side of the site (Fig. 1). One can imagine a signal peptidase specific to each pre-protein or to a tissue. However, cleavage of the signal peptide and segregation of the protein from different sources by cell-free protein synthesis systems supplemented with dog pancreatic microsomal or ascite tumor cell membranes ruled out that possibility (Bielinska et al., 1978; Blobel, Dobberstein, 1975; Lingappa et al., 1977; Lingappa et al., 1978a; Lingappa et al., 1978b).

Post-Translational Modification of Neural and Hormonal Pro-Peptides

During its biosynthesis and subsequent transport through the Golgi complex, the secretory protein is subjected to different modifications including N- or O-glycosylation, proteolysis and chemical modifications of amino acid residues such as acetylation, amidation and hydroxylation. In the past few years, much attention was focused on the mechanisms involved in the glycosylation of secretory proteins (human chorionic gonadotropin) (Bielinska, Boime 1979, Hussa 1980) and of virus protein (vesicular stomatitis virus G protein) (Kornfeld et al., 1978; Li et al., 1978; Rothman, Lodish, 1977; Tabas, Kornfeld, 1978).

The maturation of secretory proteins involves, in many cases, the proteolytic cleavage of the precursor into active substances. Examination of the primary structure of the pro-hormones and pro-proteins listed in Table I indicates that a general mechanism for their processing in active substances involves proteolytic cleavage at sequences that contain paired basic residues.

This structural requirement became first apparent in the β-LPH model proposed by Chrétien and Li (1967) and reinforced when the structure of proinsulin became known in 1968 (Chance et al., 1968) (Fig. 2). This mechanism seems to require the action of two-membrane bound proteases: (1) a trypsin-like endopeptidase cleaves specifically the polypeptide chain at the carboxyl end of the paired dibasic residues and (2) the basic amino acids are subsequently removed by a carboxypeptidase B-like exopeptidase. However, not all paired basic amino acids are cleaved. A protease with high specificity for paired basic residues has been described in the pituitary gland (Austen, 1981). The enzyme is found associated with subcellular fractions containing secretory granules indicating that proteolytic cleavage occurs late in the secretory process. Studies done on the

maturation of pro-insulin (Steiner et al., 1974) and pro-parathyroid hormone (MacGregor et al., 1976) showed that the cleavage of these pro-hormones is initiated in the Golgi complex (Kemper, 1984).

Many other enzymes have been described with trypsin-like or chymotrypsin-like activity which process pro-proteins into fragments displaying some biological activities. However, structural studies of the fragments was never done and the possibility exists that an unspecific fragment contains in part or in total the biologically active peptide. More work needs to be done on the pro-hormone maturating enzymes. We feel that we may be faced with a new set of proteolytic activities which will be different and more specific than trypsin-like and carboxypeptidase B-like described by most authors.

During biosynthesis and the subsequent intracellular transport, polypeptide chains can be subject to modifications such as disulfide bridge formation, acetylation, amidation, hydroxylation of proline and lysine residues and formation of an N-terminal pyroglutamic residue. Some of these modifications occur early during the biosynthesis of the protein (e.g. hydroxylation of collagen). Others occur only later in the Golgi complex or in the secretory granules after processing of the pro-protein by the maturation enzymes (e.g., acetylation of α-MSH). These modifications help the molecule to assume a given secondary structure and may be involved in the interaction of the molecule with its receptor.

Pro-Opiomelanocortin

1. First Biosynthesis of β-LPH, γ-LPH, and β-Endorphin in Bovine Pituitary Gland.

In 1974 to 1976, we demonstrated unequivocally that β-LPH and γ-LPH were biosynthesized in bovine pituitary glands. Soon after the isolation and the characterization of sheep and human β-endorphin (Chrétien et al., 1976; Chrétien et al., 1977), we used the highly purified sheep β-endorphin as a marker to identify its radioactive biosynthetic counterpart. Bovine pituitary slices and isolated cells of the pars intermedia incorporate radioactive methionine, leucine and lysine into substances with properties identical with those of β-endorphin (Chrétien et al., 1977; Chrétien et al., 1978a; Chrétien et al., 1978b; Chrétien et al., 1979a; Chrétien et al., 1979b; Crine et al.,1977; Seidah et al., 1979).

The material extracted from the bovine pituitary tissue was shown to contain a peptide with the same mobility as β-endorphin. Sequencing of this material showed radioactivity only in methionine at position 5. These results proved that this radioactive material is pure β-endorphin. This type of characterization was later used for the identification of rat β-LPH, γ-LPH, ACTH, α-MSH and β-endorphin (Seidah et al., 1978).

The possibility that β-endorphin is produced by degradation of β-lipotropin during the isolation process can be ruled out because Chrétien and Gilardeau, (1970) have shown that β-lipotropin is not broken down

chemically or enzymatically during the isolation process. Moreover, circular dichroism studies by St-Pierre et al. (1976) show that β-lipotropin is very stable over a wide range of pH and temperature. One can also exclude the possibility that the appearance of β-endorphin could be due to hydrolytic trypsin-like activity released from broken cells during the incubation, for three reasons: (1) protein labeling is linear with time, suggesting little lysis of pituitary cells; (2) a trypsin inhibitor was included in the incubation medium; (3) a crude pituitary extract was used as carrier, and served as substrate for any enzymatic activity which might still have been present.

2. Biosynthesis of ACTH and Related Peptides

Although the structures of ACTH and of α-MSH had been known for almost 15 years, the structural relationship between the two molecules in relation to their biosynthesis did not attract much attention until two additional facts were published: (1) that ACTH was present in high molecular weight forms (Coslovsky, Yalow, 1974; Lowry et al., 1977; Orth et al., 1970; Orth et al., 1973; Yalow et al., 1971) and (2) that CLIP was present in the pars intermedia (Scott et al., 1972; Scott et al., 1974). The most important in vitro experiments were carried out in the mid 1970's using simultaneously two experimental approaches. (Eipper and Mains, 1975, 1977, 1978a, 1978b), Eipper et al. (1976) and Mains and Eipper (1975, 1976) (Roberts and Herbert, 1977a, 1977b) and Roberts et al. (1978, 1979). Both studies confirmed the hypothesis that ACTH is biosynthesized as a large precursor form of molecular weight 28,000-31,000 which is later transformed into 13K and 4.5K molecules. Nakanishi et al. (1976, 1977a, 1977b, 1979) and Jones et al. (1977) confirmed the presence of a large precursor using mRNA from bovine pituitary glands.

Pulse-chase experiments were carried out recently showing the maturation from one molecule to the other in the rat pars intermedia (Crine et al., 1979) and in mouse At-20 tumors (Mains, Eipper 1978).

3. β-LPH and ACTH: Part of the Same Precursor

It had been known for more than a decade that β-MSH and ACTH were under similar physiological control and were also increased simultaneously in pathological conditions (Berson and Yalow, 1968). This raised the logical suspicion that both molecules came from the same cell.

Phifer et al. (1974) using immunoperoxidase staining with anti-ACTH$^{17-39}$, anti-α-MSH and anti-β-MSH, demonstrated that all these molecules were present in the so-called corticotrophs, a feature which was suspected by Dessy et al. (1973) and Furth et al. (1975), who described the β-LPH-containing cells. More recently, in collaboration with Pelletier et al. (1977) we have successfully identified β-LPH in the ACTH cells of human pituitary glands as well as in pituitaries of several other species. Our antiserum (Desranleau et al., 1972) was directed at the N-terminal portion of β-LPH and did not cross-react with β-MSH. Finally, other groups have confirmed that β-endorphin is located in the same cells (Martin et al., 1979; Weiber et al., 1978).

Although these facts had been known for some years, no attempt was made before 1976 to study the cross-reactivity of the large forms of ACTH with either a β-MSH, a β-LPH or a β-endorphin directed antiserum. Since our approach to elucidating the β-LPH maturation process was mainly to characterize chemically the secretory products, we could not do it directly since the amount of common precursor is too low. However, we were able to show that ACTH-producing tumors (AtT-20 and MtT-F4) also contain β-LPH and β-endorphin (Bourassa et al., 1978; Scherrer et al., 1978), thus completing the analogy of their secretory material with the normal pituitary gland.

Lowry et al. (1976) were the first to describe the presence of ACTH and LPH in the same molecule. Mains et al. (1977) and Roberts and Herbert (1977a, 1977b) almost simultaneously confirmed by radioimmunoassay Lowry's evidence (1976) that ACTH and LPH could be part of the same precursor. Rubinstein et al. (1978) were able to purify a small quantity of the common precursor from rat pituitaries. Crine et al. (1978) were the first to show the precursor's maturation during pulse-chase experiments, while Nakanishi et al. (1976, 1977a, 1977b, 1979) and Roberts et al. (1979) were able to obtain the final proof with their cloned cDNA sequence analysis. Such a precursor has also been found in a human ACTH ectopic tumor (Bertagna et al., 1978), in frog (Pezalla et al., 1978), and toad pituitaries (Loh, 1979).

This common precursor was first called "31K-precursor" (Eipper and Mains 1975, 1977, 1978a, 1978b; Mains and Eipper, 1976) or "precursor to ACTH, LPH and β-endorphin" (Mains and Eipper, 1977; Robert and Herbert, 1977a, 1977b) and "pro-opiocortin" (Rubinstein et al., 1978). We felt that this last piece of nomenclature was incomplete and we renamed the precursor: pro-opiomelanocortin, a name which takes into account the three main biological activities of its known secretory components (Chrétien et al., 1979a).

While continuing to work on the biosynthesis of β-LPH and its related peptides, we added the obviously important end-product ACTH. Those working on ACTH biosynthesis added β-LPH to their model. We will concentrate this section on the in vitro labeling experiments which show the maturation of the precursor during pulse-chase experiments.

A number of criteria need to be met before the existence of a biosynthetic precursor for a peptide hormone or secreted protein can be established with certainty (for similar criteria see reference Tager et al., 1975).

1. Pulse-labeling and pulse-chase experiments are essential to establish the precursor-product relationship between high and low molecular weight forms of the proteins.

2. Immunoprecipitation experiments with well-characterized antibodies raised against the hormone must precipitate the higher as well as the lower molecular weight forms of the hormones. It

could happen, however, that the high molecular weight precursor might not be recognized by the antibody to the low molecular weight molecule.

3. Peptide mapping of the precursor by means of a high resolution two-dimensional separation or by HPLC of tryptic fragments must demonstrate the existence, within the larger molecular form, of peptides characteristic of the active hormone, together with additional fragments.

4. Sequence analysis of the putative precursor either directly or through its nucleotide sequence must reveal the existence of an additional peptide covalently linked to the hormone.

Any of these findings taken alone constitutes suggestive evidence for the existence of a precursor for a given hormone, but conclusive proof can be obtained only through careful structural analysis of the putative precursor.

Gene Expression of POMC in the Human

One way to look at the expression of POMC gene in the human is to isolate and characterize the end products. We had already characterized the human β-endorphin (Dragon et al., 1977) and its sequence is shown in Fig. 3. The results obtained by Li et al. (1977) are in agreement with ours. The sequence of human β-LPH has been completed recently (Seidah et al., 1982).

The structure of human ACTH has been known for many years and all workers in the field agree on its sequence.

Up to 1980, more than 50% of the human POMC molecule was not known and this included the N-terminal prosthetic group which would contain more than 100 residues. We have recently isolated the major secretory products of this N-terminal region.

The anterior lobe of the human pituitary releases a large quantity of a 12 kilodalton glycoprotein which corresponds to the first 76 amino acids of the molecule. It has two sites of glycosylation, threonine #45 and asparagine #65. It contains 2 disulfide bridges, Cys 2-8 and Cys 20-24.

While this work was in progress, Chang et al. (1980) published the partial sequence of the genomic DNA of human POMC. These results lead to the conclusion that there has to be a peptide between the 1-76 N-terminal and the ACTH molecule.

The genomic DNA sequence predicted a peptide of 109 amino acid residues preceding the ACTH and β-lipotropin sequence (Chang et al., 1980). It was therefore proposed (Seidah and Chrétien, 1981a; Seidah et al., 1981d) that the pair of basic residues Lys_{77}-Arg_{78}, predicted from the DNA

sequence (Chang et al., 1980), are cleaved during the maturation of pro-opiomelanocortin in vivo. Therefore, a 31 amino acid "missing fragment" representing residues 79-109 must also have been generated (Chang et al., 1980; Seidah and Chrétien 1981a; Seidah et al., 1981d).

The isolation and sequence characterization of such a peptide from human pituitary extracts have been carried out by our group (Seidah et al., 1981c). It is found that the predicted DNA sequence is faithfully translated into its peptide form, but that the isolated peptide (denoted as HJP for human joining peptide) lacks the C-terminal Gly residue (Chang et al., 1980). Evidence has been presented showing that the isolated 30 residue (HJP) peptide is amidated at its COOH-terminal glutamic acid residue.

The amino acid composition and sequence of the 30 residues is identical to that expected from the nucleotide sequence (Chang et al., 1980; Cochet et al., 1982). The complete amino acid sequence of the human N-terminal glycosylated 1-76 fragment (denoted HNT) of human POMC, has been published by Takahashi et al. (1981), and Cochet et al. (1982) and the DNA sequence agrees entirely with our protein sequence. Therefore, a general maturation model is presented in Fig. 4, which takes into account our results and previously isolated peptides (Benjannet et al., 1980; Chrétien et al., 1979a; Larivière et al., 1981; Seidah and Chrétien, 1981a; Seidah et al., 1980; Seidah et al., 1981d) and the results of pulse-chase experiments (Chrétien et al., 1979a; Crine et al., 1980a; Crine et al., 1980b; Mains and Eipper, 1979; Phillips et al., 1981).

Furthermore, from the model in Fig. 4, it is seen that in the human, the maturation enzyme(s) responsible for the generation of all POMC peptides is very selective, since it cleaves preferentially at the pair of basic residues Lys-Arg (Seidah and Chrétien, 1981a; Seidah et al., 1981d). The isolation and characterization of the HJP segment add more weight to the validity of this observation, since for its production it would involve cleavage at the Lys_{77}-Arg_{78} and Lys_{110}-Arg_{111} bonds (Chang et al., 1980), followed by a so-called carboxypeptidase-B like cleavage of the C-terminal Lys-Arg pairs (Chrétien et al., 1979a; Steiner et al., 1980).

Much work needs to be done before we can accurately define the precise maturation pathway and biological functions of the various segments of this pluripotent POMC precursor molecule.

Processing of Other Precursor Molecule: The Use of Pro-Opiomelanocortin Model to Predict Their Secretion Products

With the development of DNA structure analysis, the complete sequence of the precursor can be determined either from a cDNA transcript of the mRNA or from the genomic DNA. Numerous examples have recently appeared which demonstrate the analytical power of this approach. It can be observed from such studies that the pairs of basic amino acid residues are present at most sites of cleavage (Table I). This "magic" rule became first noticed in the β/γ-LPH model proposed in 1967 (Chrétien and Li, 1967) and reinforced with pro-insulin in 1968 (Fig. 2) (Chance et al., 1968).

1. Pro-Enkephalin

If one applies this rule to the recently published structure of pro-enkephalin (Comb et al., 1982), one could predict that this multipotent gene can give rise to a number of enkephalins. As seen in Fig. 5, by assuming that the splitting is at the pair Lys-Arg as for POMC maturation, this gene gives rise to one molecule each of Met-enkephalin and Leu-enkephalin and five other longer enkephalins containing eight to 48 amino acids. If other pairs are cleaved, it could give rise to numerous other enkephalins including four molecules of Met-enkephalin, one molecule of Leu-enkephalin and two other Met-enkephalins of seven and eight amino acids each. Only pulse-chase experiments and/or purification studies will give the answer. However, all the peptides described in Fig. 5 are candidates to be secreted as such.

2. Pro-Somatostatin

Let's look now at the other models in Table I. The first somatostatin isolated and characterized by Brazeau et al. (1973) has 14 amino acids. A somatostatin 28 amino acids long only has been found so far by Pradayrol et al. (1980). The secretion of this new larger somatostatin indicates a cleavage at one arginyl.

3. Pro-Insulin

In pro-insulin (Table I and Fig. 2), there is a release of the C-peptide which has yet to find a biological role.

4. Pro-Parathormone

The PTH model (Table I) (MacGregor et al., 1976) (see also Kemper, 1984) is the simplest of all since the active component is at the C-terminus while it has a very short pro-segment (5 amino acid residues).

5. Pro-Vasopressin (pro-pressophysin)

The vasopressin precursor, also called pro-pressophysin, has also been recently sequenced (Land et al., 1982). It is interesting to note that a glycopeptide we have recently isolated (Seidah et al., 1981b) and proposed it to be part of the vasopressin precursor, is the C-terminal fragment of the precursor with vasopressin at the C-terminus and neurophysin in the centre (Table I).

6. Pro-Calcitonin

Finally, the calcitonin model (Table I) is more standard with calcitonin in the middle between pairs of Lys-Arg and a C-terminus fragment of unknown biological significance. In addition, the calcitonin gene is expressed in the brain, but then gives us a different set of end-products, one of which, called CGRP, might have a biological function (Rosenfeld et al., 1983).

Conclusion

Through the knowledge of the chemistry of pro-hormones, one can envisage their maturation processes during their transport from the endoplasmic reticulum to the secretion granules. More important, it is possible to foresee the biosynthesis of new peptides whose existence is not suspected by the regular isolation methods which have been the only accessible methods so far.

TABLE 1. SEQUENCES AT MAIN CLEAVAGE SITES IN PRECURSOR PROTEINS

LEGEND

E Glutamic Acid
F Phenylalanine
G Glycine
K Lysine
L Leucine
R Arginine
V Valine

Met-enkephalin

Leu-enkephalin

MSH's

FIGURE 1

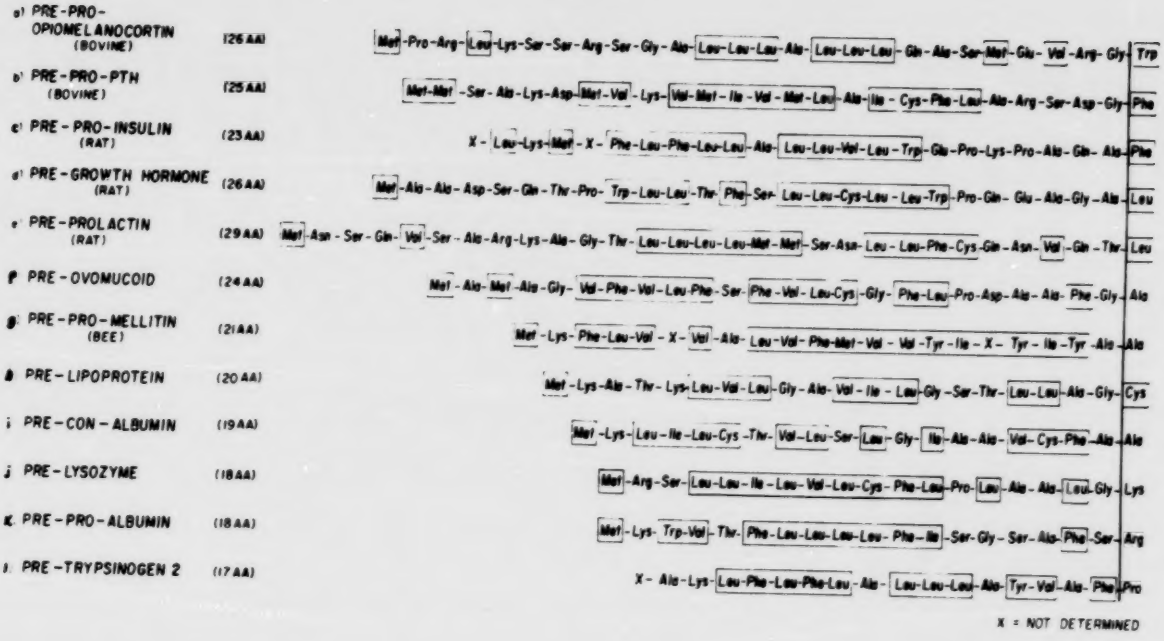

a) PRE-PRO-OPIOMELANOCORTIN (BOVINE) (126 AA)
b) PRE-PRO-PTH (BOVINE) (25 AA)
c) PRE-PRO-INSULIN (RAT) (23 AA)
d) PRE-GROWTH HORMONE (RAT) (26 AA)
e) PRE-PROLACTIN (RAT) (29 AA)
f) PRE-OVOMUCOID (24 AA)
g) PRE-PRO-MELLITIN (BEE) (21 AA)
h) PRE-LIPOPROTEIN (20 AA)
i) PRE-CON-ALBUMIN (19 AA)
j) PRE-LYSOZYME (18 AA)
k) PRE-PRO-ALBUMIN (18 AA)
l) PRE-TRYPSINOGEN 2 (17 AA)
X = NOT DETERMINED

FIGURE 2

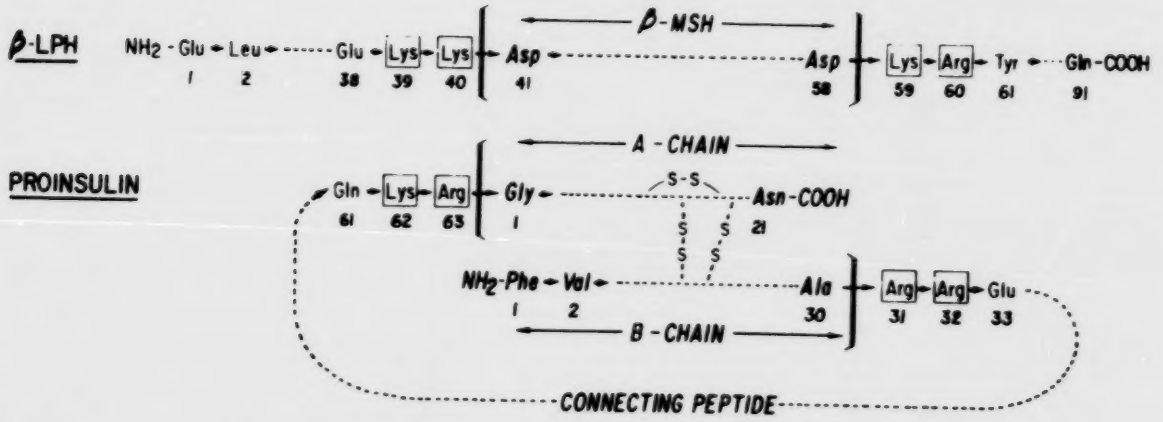


FIGURE 3

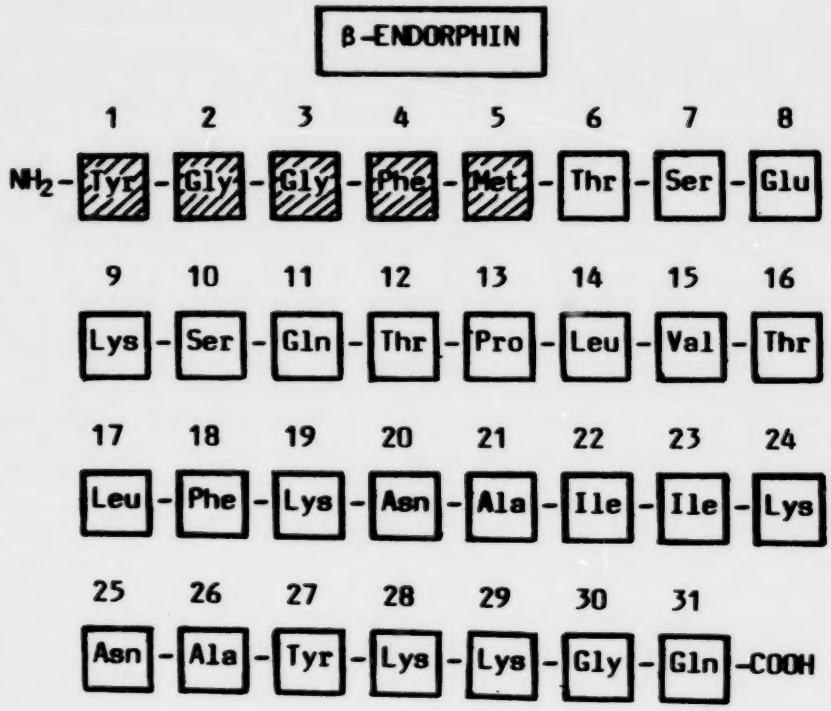

β-ENDORPHIN
1 2 3 4 5 6 7 8
NH2-Tyr-Gly-Gly-Phe-Met-Thr-Ser-Glu
9 10 11 12 13 14 15 16
Lys-Ser-Gln-Thr-Pro-Leu-Val-Thr
17 18 19 20 21 22 23 24
Leu-Phe-Lys-Asn-Ala-Ile-Ile-Lys
25 26 27 28 29 30 31
Asn-Ala-Tyr-Lys-Lys-Gly-Gln-COOH

FIGURE 4

HUMAN PITUITARY PRO-OPIOMELANOCORTIN AND ITS KNOWN MATURATION PRODUCTS

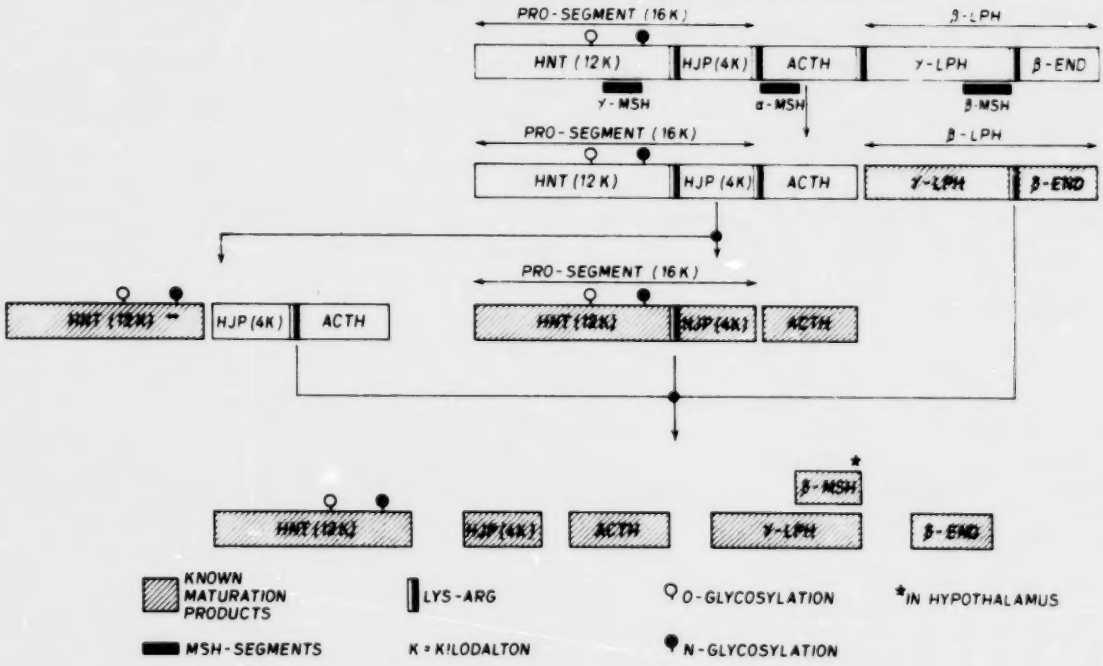
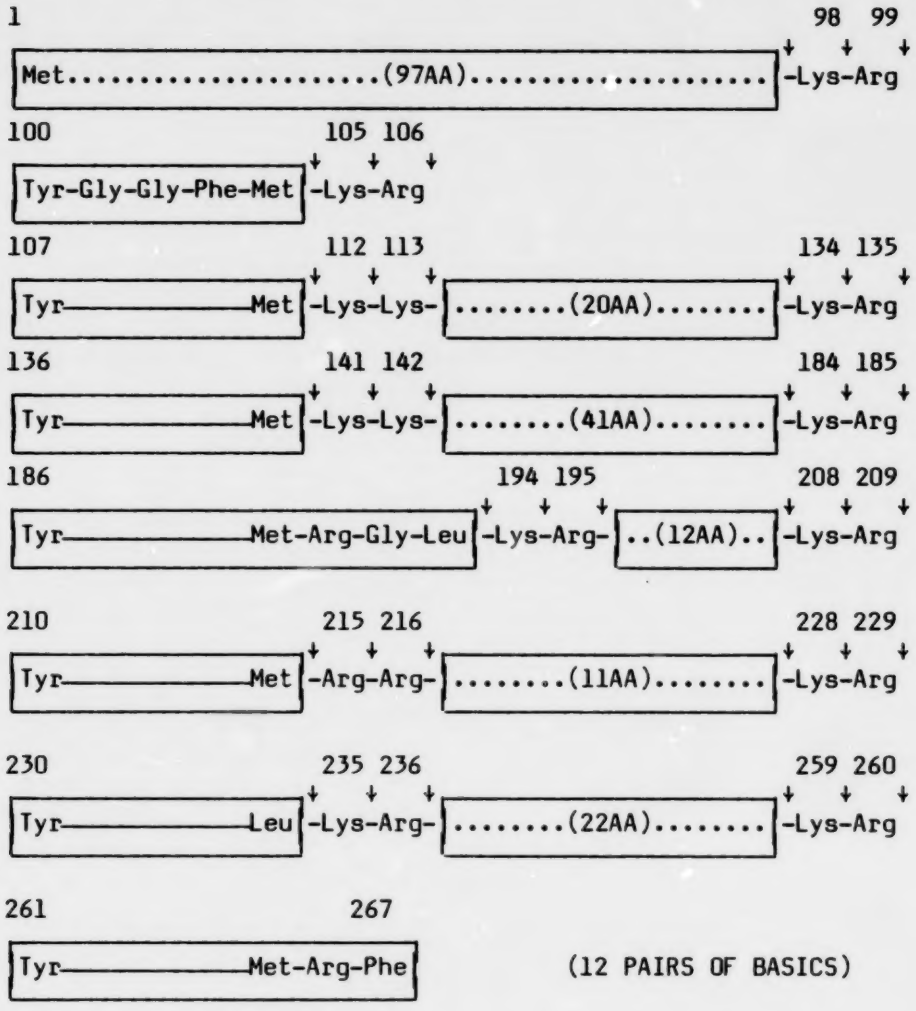


FIGURE 5

HUMAN PRE-PRO-ENKEPHALIN



COMB AND HERBERT (NATURE 295:663, 1982)

LEGENDS TO FIGURES

Fig. 1. Amino acid sequence of the "signal peptides" of some secretory proteins.

Fig. 2. Amino acid sequences of β-LPH and pro-insulin at their site of enzymatic cleavage.

Fig. 3. Sequence of human β-endorphin.

Fig. 4. General model for maturation of human POMC molecule. The hatched areas represent isolated and completely characterized peptides from human pituitaries. The molecular weights of the human N-terminal (HNT) 76 residues glycosylated peptide and the human joining peptide (HJP) isolated are given in parenthesis as 12 and 4 kilodaltons (K) respectively. The glycosylation sites of HNT are at Thr_{45} and Asn_{65}. Exclusive cleavage at the pair of basic residues Lys-Arg is emphasized, together with the position of the three MSH segments within the POMC precursor. The possibility of generating an HJP peptide attached to HNT or to ACTH is proposed, based on the low yields of the HJP peptide isolated.

Fig. 5. The complete sequence of pre-pro-enkephalin as determined from the cDNA structure by Comb et al. (1982). The sites of possible cleavages with the description of the end-products. If the Lys-Arg pair is the only one cleavage, the pro-molecule will give rise to one molecule each of Met and Leu- enkephalin, one Met-enkephalin of 27 residues (amino acid 107 to 133), one Met-enkephalin of 48 residues (amino acid 136 to 183), one Met-enkephalin of 8 residues (amino acid (186 to 193), one Met-enkephalin of 18 residues (amino acid 210 to 228), one Met-enkephalin of seven residues (amino acid 261 to 267), three non-enkephalin related peptides, one of 97 residues (minus the signal peptide), one of 12 residues (amino acid 196 to 207) and one of 22 residues (amino acid 237 to 258). If the other pairs are cleaved, the secretory pattern becomes more simple. One gets one molecule of Leu-enkephalin, three types of Met-enkephalin, four molecules of regular Met-enkephalin, one each of seven and eight amino acids.

REFERENCES

Austen, B.M. (1981) Isolation from pituitary granules and properties of a protease specific for consecutive basic residues. In: Peptides, Synthesis - Structure - Function, (D.H. Rich & E. Gross, eds.) Pierce Chemical Co., Rockford, Ill. pp. 493-496.

Benjannet, S., Seidah, N.G., Routhier, R., and Chrétien, M. (1980) A novel human pituitary peptide containing the γ-MSH sequence. Nature 285:415-416.

Berson, S.A., and Yalow, R.S. (1968) Peptide hormones in plasma. Harvey Lectures 62:107-163.

Bertagna, X., Nicholson, W.E., Sorenson, G.D., Pettengill, O.S., Mount, C.D., and Orth, D. (1978) Corticotropin, lipotropin, and β-endorphin production by a human non pituitary tumor in culture: Evidence for a common precursor. Proc. Natl. Acad. Sci. USA 75:5160-5164.

Bielinska, M., and Boime, I. (1979) Glycosylation of human chorionic gonadotropin in mRNA-dependent cell-free extracts: post-translational processing of an asparagine-linked mannose-rich oligosaccharide. Proc. Natl. Acad. Sci. USA 76:1208-1212.

Bielinska, M., Grant, G., and Boime, I. (1978) Processing of placental peptide hormones synthesized in lysates containing membranes derived from tunicamycin-treated ascites tumor cells. J. Biol. Chem. 253:7117-7119.

Blobel, G., and Dobberstein, B. (1975) Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. J. Cell. Biol. 67:835-851.

Bourassa, M., Scherrer, H., Pezalla, P.D., Lis, M. and Chrétien, M. (1978) Immunoreactive β-endorphin in the rat mammotropic transplantable tumor (MtT-F4). Cancer Res. 38:1568-1571.

Brazeau, P., Vale, W., Burgus, R., Ling, N., Butcher, M., Rivier, J., and Guillemin, R. (1973) Hypothalamic polypeptide that inhibits the secretion of immunoreactive pituitary growth hormone. Science 179:77-79.

Chance, R.E., Ellis, R.M., and Bromer, W.W. (1968) Porcine pro-insulin: characterization and amino acid sequence. Science 161:165-167.

Chang, A.C.Y., Cochet, M., and Cohen, S.N. (1980) Structural organization of human genomic DNA encoding and pro-opiomelanocortin peptide. Proc. Natl. Acad. Sci. USA 77:4890-4894.

Chrétien, M., and Gilardeau, C. (1970) Effects of glacial acetic acid on the extraction, chemical stability, and biological activities of β-LPH. Can. J. Biochem. 48:511-516.

Chrétien, M., and Li, C.H. (1967) Isolation, purification and characterization of γ-lipotropic hormone from sheep pituitary glands. Can. J. Biochem. 45:1163-1174.

Chrétien, M., Benjannet, S., Dragon, N., Seidah, N.G., and Lis, M. (1976) Isolation of peptides with opiate activity from sheep and human pituitaries: relationship to β-lipotropin. Biochem. Biophys. Res. Commun. 72:472-478.

Chrétien, M., Seidah, N.G., Benjannet, S., Dragon, N., Routhier, R., Motomatsu, T., Crine, P., and Lis, M. (1977) A β-LPH precursor model: recent developments concerning morphine-like substances. Ann. N.Y. Acad. Sci. 297:84-107.

Chrétien, M., Crine, P., Lis, M., and Seidah, N.G. (1978a) Biosynthesis of β-endorphin. Bull. Schweiz. Akad. Med. Wiss 34:155-169.

Chrétien, M., Crine, P., Lis, M., Gianoulakis, C., Gossard, F., Benjannet, S., and Seidah, N.G. (1978b) Biosynthesis of β-endorphin from β-lipotropin and a large precursor molecule. In: Characteristics and Functions of Opioids. Developments in Neurosciences", Vol. IV, pp. 345-357, Elsevier-North Holland, New York.

Chrétien, M., Benjannet, S., Gossard, F., Gianoulakis, C., Crine, P., Lis, M., and Seidah, N.G. (1979a) From β-lipotropin to β-endorphin and pro-opiome-lanocortin. Can. J. Biochem. 57:1111-1121.

Chrétien, M., Crine, P., Lis, M., Benjannet, S., and Seidah, N.G. (1979b) β-lipotropin: Pro-endorphin. In "Central Nervous System Effects of Hypothalamic Hormones and Other Peptides", pp. 237-251, Ravan Press, New York.

Cochet, M., Chang, A.C.Y., and Cohen, S.N. (1982) Characterization of the structural gene and putative 5'-regulatory sequences for human pro-opiomelanocortin. Nature 297:35-338.

Comb, M., Seeburg, P.H., Adelman, J., Eiden, L., and Herbert, E. (1982) Primary structure of the human Met- and Leu-enkephalin precursor and its mRNA. Nature 295:663-666.

Coslovsky, R., and Yalow, R.S. (1974) Influence of the hormonal forms of ACTH on the pattern of corticosteroid secretion. Biochem. Biophys. Res. Comm. 60:1351-1356.

Crine, P., Benjannet, S., Seidah, N.G., Lis, M. and Chrétien, M. (1977) In virto biosynthesis of β-endorphin in pituitary glands. Proc. Natl. Acad. Sci. USA 74:1403-1406.

Crine, P., Gianoulakis, C., Seidah, N.G., Gossard, F., Pezalla, P.D., Lis, M., and Chrétien, M. (1978) Biosynthesis of β-endorphin from β-lipotropin and a larger molecular weight precursor in rat pars intermedia. Proc. Nat. Acad. Sci. USA 75:4719-4723.

Crine, P., Seidah, N.G., Routhier, R., Gossard, F. and Chrétien, M. (1980a) Processing of two forms of the common precursor to α-melanotropin and β-endorphin in the rat pars intermedia. Eur. J. Biochem. 110:387-396.

Crine, P., Seidah, N.G., Jeannotte, L., and Chrétien, M. (1980b) Two large glycoprotein fragments related to the NH_2-terminal part of the adrenocor-ticotropin-β-lipotropin precursor are the end-products of the maturation process in the rat pars intermedia. Can. J. Biochem. 58:1318-1322.

Desranleau, R., Gilardeau, C., and Chrétien, M. (1972) Radioimmunoassay of ovine β-lipotropic hormone. Endocrinology 91:1004-1010.

Dessy, C., Herlant, M., and Chrétien, M. (173) Détection par immuno-fluorescence des cellules synthétisant la lipotropin. C.R. Acad. Sci. Paris 276:335-338, 1973.

Dragon, N., Seidah, N.G., Lis, M., Routhier, R., and Chrétien, M. (1977) Primary structure and orphine-like activity of human β-endorphin. Can. J. Biochem. 55:666-670.

Eipper, B.A., and mains, R.E. (1975) High molecular weight forms of adrenocorticotropic hormone in the mouse pituitary and in a mouse pituitary tumor cell line. Biochemistry 14:3836-3844.

Eipper, B.A., and Mains, R.E. (1977) Peptide analysis of a glycoprotein form of adrenocorticotropic hormone. J. Biol. Chem. 24:8821-8832.

Eipper, B.A., and Mains, R.E. (1978a) Analysis of the common precursor to corticotropin and endorphin. J. Biol. Chem. 253:5732-5744.

Eipper, B.A., and Mains, R.E. (1978b) Existence of a common precursor to ACTH and endorphin in the anterior and intermediate lobes of the rat pituitary. J. Supramol. Struct. 8:247-262.

Eipper, B.A., Mains, R.E., and Guenzi, D. (1976) High molecular weight forms of adrenocorticotropic hormone are glycoproteins. J. Biol. Chem. 251:4121-4126.

Furth, J., Chrétien, M., Lis, M., Bélanger, A., Moy, P., and Grauman, J. (1975) Multipotent lipotropic hormones. in search of a pituitary cell producing multipotent LPH. Arch. Pathol. 99:572-581.

Hussa, R.O. (1980) Biosynthesis of human chorionic gonadotropin. Endocrine Reviews 1:268-294.

Jacobs, J.W., Goodman, R.H., Chin, W.W., Dee, P.C., Habener, J.F., Bell, N.H., and Potts, J.T. Jr. (1981) Calcitonin messenger RNA encodes multiple polypeptides in a single precursor. Science 213:457-459.

Jackson, R.C., and Blobel, G. (1980) Post-translational processing of full-length presecretory proteins with canine pancreatic signal peptidase. Ann. N.Y. Acad. Sci. 343:391-403.

Jones, R.E., Pulkrabek, P., and Grunberger, D. (1977) Mouse pituitary tumor mRNA directed cell-free synthesis of polypeptides that are cross-reactive with adrenocorticotropic hormone antiserum. Biochem. Biphys. Res. Comm. 74:1490-1495.

Kember, B. (1984) Biosynthesis and secretion of parathyroid hormone. Cell Biology of the Secretory Process, pp. 443-480.

Kornfeld, S., Li, E., and Tabas, I. (1978) The synthesis of complex-type oligosaccharides. II. Characterization of the processing intermediates in the synthesis of the complex oligosaccharide units of the vesicular stomatitis virus G protein. J. Biol. Chem. 253:7771-7778.

Land, H., Schütz, G., Schmale, H., and Richter, D. (1982) Nucleotide sequence of cloned cDNA encoding bovine arginine vasopressin-neurophysin II precursor. nature 295:299-303.

Larivière, N., Seidah, N.G., and Chrétien, M. (1981) Complete sequence of the glycosylated amino-terminal segment of porcine pro-opiomelanocortin. Int. J. Pept. Prot. Res. 18:487-491.

Li, C.H., Yamashiro, D., Chung, D., and Doneen, B.A. (1977) Isolation, structure, synthesis and morphine-like activity of β-endorphin from human pituitary glands. Ann. N.Y. Acad. Sci. 297:158-166.

Li, E., Tabas, I., and Kornfeld, S. (1978) The synthesis of complex-type oligosaccharides. I. Structure of the lipid-linked oligosaccaride precursor of the complex-type oligosaccharides of the vesicular stomatitis virus G protein. J. Biol. Chem. 253:7762-7770.

Lingappa, V.R., Devillers-Thiery, A., and Blobel, D. (1977) Nascent pre-hormones are intermediates in the biosynthesis of authentic bovine pituitary growth hormone and prolactin. Proc. Nat. Acad. Sci. USA 74:2432-2436.

Lingappa, V.R., Katz, F.N., Loddish, H.F., and Blobel, G. (1978a) A signal sequence for the insertion of a transmembrane glycoprotein. Similarities to the signals of secretory proteins in primary structure and function. J. Biol. Chem. 253:8867-8870.

Lingappa, V.R., Prusad, R., Ebner, K., and Blobel, G. (1978b) Coupled cell-free synthesis, segregation, and core glycosylation of a secretory protein. Proc. Nat. Acad. Sci. USA 75:2338-2342.

Loh, Y.P. (1979) Immunological evidence for two common precursors to corticotropins, endorphins, and melanotropin in the neurointermediate lobe of the toad pituitary. Proc. Nat. Acad. Sci. USA. 76:796-800.

Lowry, P.J., Hope, J., and Silman, R.E. (1976) The evolution of corticotropin-melanotropin and lipotropins. Excerpt. Med. Int. Congr. Ser. 402:71-76.

Lowry, P.J., Silman, R.E., Hope, J., and Scott, A.P. (1977) Structure and biosynthesis of peptides related to corticotropins and β-melanotropins. Ann. N.Y. Acad. Sci. 297:49-60.

Lund, P.K., Goodman, R.H., Dee, P.C. and Habener, J.F. (1982) Pancreatic preproglycagon cDNA contains two glucagon-related coding sequences arranged in tandem. Proc. Natl. Acad. Sci. USA 79:345-349.

MacGregor, R.R., Chu, L.L.H., and Lohn, D.V. (1976) Conversion of proparathyroid hormone to parathyroid hormone by a particulate enzyme of the parathyroid gland. J. Biol. Chem. 251-6711-6716.

Mains, R.E., and Eipper, B.A. (1975) Molecular weights of adrenocorticotropic hormone in extracts of anterior and intermediate-posterior lobes of mouse pituitary. Proc. Nat. Acad. Sci. USA 72:3565-3569.

Mains, R.E., and Eipper, B.A. (1976) Biosynthesis of adrenocorticotropic hormone in mouse pituitary tumor cells. J. Biol. Chem. 251:4115-4120.

Mains, R.E., and Eipper, B.A. (1978) Coordinate synthesis of corticotropins and endorphins by mouse pituitary tumor cells. J. Biol. Chem. 253:651-655.

Mains, R.E., and Eipper, B.A. (1979) Synthesis and secretion of corticotropins, melanotropins, and endorphins by rat intermediate pituitary cells. J. Biol. chem. 254:7885-7894.

Mains, R.E., Eipper, B.A., and Ling, N. (1977) Common precursor to corticotropins and endorphins. Proc. Nat. Acad. Sci. USA 74:3014-3018.

Martin, R., Weber, E., and Voigt, K.H. (1979) Localization of corticotropin- and endorphin-related peptides in the intermediate lobe of the rat pituitary. Cell. Tissue Res. 196:307-319.

Nakanishi, S., Taii, S., Hirata, Y., Matsukura, S., Imura, H., and Numa, S. (1976) A large product of cell-free translation of messenger RNA coding for corticotropin. Proc. Nat. Acad. Sci. USA 73:4319-4323.

Nakanishi, S., Inoue, A., Taii, S., and Numa, S. (1977a) Cell-free translation product containing corticotropin and β-endorphin encoded by messenger RNA from anterior lobe and intermediate lobe of bovine pituitary. FEBS Lett. 84:105-109.

Nakanishi, S., Kita, T., Taii, S., Imura, H., and Numa, S. (1977b) Glucocorticoid effect on the level of corticotropin messenger RNA activity in rat pituitary. Proc. Nat. Acad. Sci. USA 74:3283-3286.

Nakanishi, S., Inoue, A., Kita, T., Nakamura, M., Chang A.C.Y., Cohen, S.N., and Numa, S. (1979) Nucleotide sequence of cloned cDNA for bovine corticotropin-β-lipotropin precursor. Nature 278:432-427.

Orth, D.N., Nicholson, W.E., Shapiro, M., and Byyny, R. (1970) Andrenocorticotropic hormone (ACTH) and melanocyte-stimulating hormone (MSH) production by a single cell. Endocrine Soc. 52nd Meeting, Abst. #140, St. Louis, Mo.

Orth, D.N., Nicholson, W.E., Mitchell, W.M., Island, D.P., Shipiro, M., and Byyny, R.L. (1973) ACTH and MSH production by a single cloned mouse pituitary tumor cell line. Endocrinology 92:385-393.

Pelletier, G., Leclerc, R., Labrie, F., Côté, J., Chrétien, M., and Lis, M. (1977) Immunohistochemical localization of β-lipotropic hormone in the pituitary gland. Endocrinology 100:770-776.

Pezalla, P.D., Seidah, N.G., Benjannet, S., Crine, P., Lis, M., and Chrétien, M. (1978) Biosynthesis of β-endorphin, β-lipotropin and the putative ACTH-LPH precursor in the frog pars intermedia. Life Sci. 23:281-2291.

Phifer, R.F., Orth, D.N., and Spicer, S.S. (1974) Specific demonstration of the human hypophyseal adrenocortico-melanotropic (ACTH/MSH) cell. J. Clin. Endo. Met. 39:684:692.

Phillips, M.A., Budarf, M.L., and Herbert, E. (1981) Glycosylation events in the processing and secretion of pro ACTH-endorphin in mouse pituitary tumor cells. Biochemistry 20:1666-1675.

Pradayrol, L., Jornvall, H., Mutt, V., and Ribet, A. (1980) N-terminally extended somatostatin: the primary structure of somatostatin-28. FEBS Let. 109:55-58.

Roberts, J.L., and Herbert, E. (1977a) Characterization of a common precursor to corticotropin and β-lipotropin: cell-free synthesis of the precursor and identification of corticotropin peptides in the molecule. Proc. Nat. Acad. Sci. USA 74:4826-4830.

Roberts, J.L., and Herbert, E. (1977b) Characterization of a common precursor to corticotropin and β-lipotropin peptides and their arrangement relative to corticotropin in the precursor synthesized in a cell-free system. Proc. Nat. Acad. Sci. USA 74:5300-5304.

Roberts, J.L., Phillips, M., Rosa, P.A., and Herbert, E. (1978) Steps involved in the processing of common precursor forms of adrenocorticotropin and endorphin in cultures of mouse pituitary cells. Biochemistry 17:3609-3618.

Roberts, J.L., Seeburg, P.H., Shine, J., Herbert, E., Baxter, J.D., and Goodman, H.M. (1979) Corticotropin and β-endorphin: construction and analysis of recombinant DNA complementary to mRNA for the common precursor. Proc. Nat. Acad. Sci. USA 76:2153-2157.

Rosenfeld, M.G., Mermod, J.-J., Amara, S.G., Swanson, L.W., Sawchenko, P.E., Rivier, J., Vale, W.W., and Evans R.M. (1983) Production of a novel neuropeptide encoded by the calcitonin gene via tissue-specific RNA processing. Nature 304:129-135.

Rothman, J.E., and Lodish, H.F. (1977) Synchronised transmembrane insertion and glycosylation of a nascent membrane protein. Nature 269:775-779.

Rubinstein, M., Stein, S., and Udenfriend, S. (1978) Characterization of pro-opiocortin, a precursor to opioid peptides and corticotropin. Proc. Nat. Acad. Sci. USA 75:669-671.

St-Pierre, S., Gilardeau, C., nd Chrétien, M. (1976) Circular dichroism studies of sheep β-lipotropic hormone. Can. J. Biochem. 54:992-998.

Scherrer, H., Seidah, N.G., Benjannet, S., Lis, M., Pezalla, P.D., Bourassa, M., and Chrétien, M. (1978) β-endorphin and β-lipotropin secretion by an ACTH-secreting mouse pituitary tumor. FEBS Lett. 90:353-356.

Scott, A.P., Bennett, H.P.J., Lowry, P.J., McMartin, C., and Ratcliffe, J.G. (1972) corticotrophin-like intermediate lobe peptide (CLIP)- a new pituitary and tumor peptide. J. Endocrinol. 55:36-37.

Scott, A.P., Lowry, P.J., Bennett, H.P.J., McMartin, C., and Ratcliffe, J.G. (1974) Purification and characterization of porcine corticotropin-like intermediate lobe peptide. J. Endocrinol. 61:369-380.

Seidah, N.G., and Chrétien, M. (1981a) Complete amino acid sequence of human pituitary glycopeptide: an important maturation product of pro-opiomelanocortin. Proc. Natl. Acad. Sci. USA 78:4236-4240.

Seidah, N.G., Gianoulakis, C., Crine, P., Lis, M., Benjannet, S., Routhier, R., and Chrétien, M., (1978) In vitro biosynthesis and chemical characterization of β-lipotropin, γ-lipotropin, and β-endorphin in rat pars intermedia. Proc. Nat. Acad. Sci. USA 75:3153-3157.

Seidah, N.G., Chan, J., Mardini, G., Benjannet, S., Chrétien, M., Boucher, R., and Genest, J. (1979) Specific cleavage of β-LPH and ACTH by tonin: release of an opiate-like peptide, β-LPH (61-78). Biochem. Biophys. Res. Comm. 86:1002-1013.

Seidah, N.G., Benjannet, S. and Chrétien, M. (1981b) The complete sequence of a novel human pituitary glycopeptide homologous to pig posterior pituitary glycopeptide. Biochem. Biphys. Res. Commun. 100:901-907.

Seidah, N.G., Benjannet, S., Routhier, R., De Serres, G., Rochemont, J., Lis, M., and Chrétien, M. (1980) Purification and characterization of the N-terminal fragment of pro-opiomelanocortin from human pituitaries: homology to the bovine sequence. Biochem. Biophys. Res. Commun. 95:1417-1424.

Seidah, N.G., Rochemont, J., Hamelin, J., Benjannet, S., and Chrétien, M. (1981c) The missing fragment of the pro-sequence of human pro-opiomelanocortin: sequence and evidence for C-terminal anmidation. Biochem. Biophys. Res. Commun. 102:710-716.

Seidah, N.G., Rochemont, J., Hamelin, J., Lis, M., and Chrétien, M. (1981d) Primary structure of the major human pituitary glycopeptide pro-opiomel-anocortin NH_2-terminal glycopeptide: Evidence for an aldosterone-stimulating activity. J. Biol. Chem. 256:7977-7984.

Seidah, N.G., Hsi, K.L., Chrétien, M., Barat, E., Patthy, A., and Graf, L. (1982) The primary structure of human β-lipotropin: Further peptide sequencing resolves the controversy and suggests the existence of only one human β-lipotropin. FEBS Letters 147:267-272.

Steiner, D.F., Kemmler, W., Tager, H.S., and Peterson, J.D. (1974) Proteolytic processing in the biosynthesis of insulin and other proteins. Fed. Proc. 33:2105-2115.

Steiner, D.F., Quinn, P.S., Chan, S.J., Marsh, J., and Tager, H.S. (1980) Processing mechanisms in the biosynthesis of proteins. Ann. N.Y. Acad. Sci. 343:1-16.

Tabas, I., and Kornfeld, S. (1978) The synthesis of complex-type oligosaccharides. III. Identification of an α-d-mannosidase activity involved in a late stage of processing of complex-type oligosaccharides. J. Biol. Chem. 253:7779-7786.

Tager, H.S., Rubenstein, A.H., and Steiner, D.F. (1975) Methods for the assessment of peptide precursors. Studies on insulin biosynthesis. In: "Methods in Enzymology", Vol. XXXVII, pp. 326-345, Academic Press, New York.

Takahashi, H., Teranishi, Y., Nakanishi, S., and Numa, S. (1981) Isolation and structural organization of the human corticotropin-β-lipotropin precursor gene. FEBS Lett. 135:97-102.

Weber, E., Voigt, K.H., and Martin, R. (1978) Concommitant storage of ACTH- and endorphin-like immunoreactivity in the secretory granules of anterior pituitary corticotrophs. Brain Res. 157:385-390.

Weissbach, H. (1979) In: "Ribosomes: Structure, Function and Genetics" (G. Chambliss, G.R. Craven, J. Davies, K. Davis, L. Kahan, and M. Nomura, eds.) University Park Press, pp. 377-411, Baltimore, MD.

Yalow, R.D., and Berson, S.A. (1971) Size heterogeneity of immunoreactive human ACTH in plasma and in extracts of pituitary glands and ACTH-producing thymoma. Biochem. Biphys. Res. Comm. 44:439-445.

Zwizinski, C., and Wickner, W. (1980) Purification and characterization of leader (signal) peptidase from Escherichia coli. J. Biol. Chem. 255:7973-7977.

RADIOIMMUNOASSAYS

ADVANCES IN THE DEVELOPMENT OF RADIOIMMUNOASSAYS TOWARDS DRUGS USED IN THE TREATMENT OF MENTAL ILLNESS

G. MCKAY, Ph.D.
K.K. MIDHA, Ph.D.

ABSTRACT

Many of the psychotropic drugs currently in wide use today are extremely difficult to measure in the plasma of patients receiving these agents. Some of the difficulties in monitoring therapeutic levels of these drugs by conventional methods can be attributed to the low doses given, their extensive binding, metabolism to numerous metabolites and inherent chemical instability displayed by these agents. Recent advances in the area of radioimmunoassay (RIA) and more specifically drug-specific RIA have allowed, in some cases for the first time, the quantitative prowess needed to monitor plasma levels of some of these drugs and their metabolites. The advantages, disadvantages and approach to the development of RIA procedures is illustrated using specific examples of RIA procedures for chlorpromazine and trifluoperazine and their major metabolites. The utility and application of monoclonal antibodies and stereoselective RIA techniques are also discussed.

The bioanalysis of many of the clinically important neuroleptic drugs used in the treatment of patients with psychotic disorders has been, and continues to be, extremely challenging analytically. Many of the difficulties encountered during such analyses can be envisaged if one considers the analysis of the phenothiazine antipsychotics used in the treatment of schizophrenia. The introduction of phenothiazine drug therapy in the treatment of schizophrenia has allowed many of the individuals suffering from the disease to become de-institutionalized and return to normal society. However, such treatment has not been without adverse effects for the patients involved. Patients treated over a long period with these neuroleptics develop motor side effects, such as tardive dyskinesia. This side effect is apparently dose related and, thus, patients should be maintained on the lowest dose required for therapeutic efficacy. Simultaneously, the blood levels required to elicit a therapeutic effect might be above a certain threshold level. Further factors complicating all of this are the marked intra- and interindividual differences in the blood levels seen at any given dose as well as individual differences in the levels at which the therapeutic effect is displayed. In attempting to find plasma concentration versus clinical response correlates, some investigators have shown the strongest correlation with clinical improvement when the pharmacologically active drug and/or metabolite is considered in relation to the pharmacologically inactive metabolites. This has been the case with chlorpromazine (CPZ) where patients, judged to be under good control, had

relatively higher levels of biologically active 7-hydroxyCPZ metabolite than patients who were poorly controlled and in whom the biologically inactive CPZ sulfoxide metabolite predominated (MacKay et al., 1974). Thus, the importance of monitoring pharmacologically important metabolites, such as the sulfoxide, may also aid in the understanding of the therapeutic efficacy of these agents.

As previously stated, the analyses of these drugs are not without difficulties. In fact, such analyses are hampered by virtue of the facts that they are extensively metabolized to numerous metabolites, administered in relatively low doses, highly protein bound and unstable at all stages of sample preparation. Consequently, procedures which require extraction and/or derivatization may create artifactual problems. Recently, it has been observed that the N-oxide metabolite can be reconverted to the parent compound during conventional extraction procedures in the analysis of CPZ (Hubbard et al., 1985; Hawes et al., 1985).

Recently, procedures based on immunoassay have been developed to avoid many of these problems. Specifically, radioimmunoassay (RIA) procedures have been successful in the plasma analysis of many drugs, including psychotropics. RIA is a biological technique which allows the quantitation of a variety of compounds present in biological media based on the competitive immune response of an antibody towards both a radiolabelled and non-labelled antigen. The focus of this article is to address the utility of RIA in the quantitation of drugs used in the treatment of various mental disorders. This will best be carried out by first briefly discussing some of the basic requirements of RIA followed by an outline of the approach taken by these investigators with specific examples in RIA of various psychotropic drugs and their metabolites.

Basic Requirements of RIA

The RIA technique itself is based on a competitive protein binding assay as mentioned earlier and can be illustrated as follows:

$$\text{Drug} + \text{Antibody} + \text{Drug}^* \rightleftharpoons \underbrace{\text{Antibody-Drug} + \text{Antibody-Drug}^*}_{\text{Bound}} + \underbrace{\text{Drug} + \text{Drug}^*}_{\text{Free}}$$

In this particular example, the compound being measured is a drug. The utilization of RIA for the measurement of drugs was first reported in 1968 by Oliver for the drug digitoxin (Oliver et al., 1968). The motivating forces behind the utilization of RIA for drug analysis are borne out if one considers the advantages versus disadvantages of RIA over other techniques used in drug analysis.

In terms of advantages, RIA procedures demonstrate high to ultrasensitivity, thereby enabling the quantitation of subnanogram to nanogram amounts of analyte. This is especially advantageous in the analysis of the

*indicates radiolabelled drug molecule.

more potent antipsychotic drugs such as trifluoperazine (TFP) or fluphenazine where peak plasma concentrations after a single oral dose are only 1-2 ng/ml (Midha et al., 1983, 1983a; Midha et al., 1981). Also, in those cases where quantitation of a trace metabolite is desirable, the ultrasensitivity of RIA is advantageous. Furthermore, only small sample sizes are required and sample preparation is usually minimal without the need for derivatization or extraction. In addition, large numbers of samples can be analyzed with a minimal turnaround time and because the technique is relatively simple, there is no need for highly skilled personnel or expensive instrumentation such as with other ultrasensitive techniques as GC/MS. Another consideration of RIA is that the method may or may not demonstrate specificity. In terms of advantages, RIA procedures have been developed which demonstrate stereoselectivity for specific isomers such as in the case of ephedrine (Midha et al., 1983b) and doxepin (Midha et al., unpublished observations).

In terms of disadvantages, RIA procedures require months of preparatory work in order to obtain antibodies which may or may not be useful. There is always a certain degree of non-specific binding which may or may not make the technique useful. Finally, any RIA procedure developed should always be confirmed through the development of a chemical method which may require further investment of time and money.

Clearly the advantages of RIA outweigh the disadvantages as evidenced by the ever-increasing number of new RIA procedures which are appearing. Obviously a key requirement for the development of any RIA procedure is the production of a drug-specific antibody. This requirement is compounded by the fact that most drugs by themselves are not antigenic. Thus, the initial stages of an RIA procedure for drug analysis requires the synthesis of an antigen. This difficulty can be overcome by covalently linking the non-antigenic molecule to a macromolecule such as bovine serum albumin (BSA). Thus, dependent on the drug in question, a preliminary synthesis of a drug hapten may be necessitated in order to either covalently link the drug to the protein or in order to adequately expose the drug away from the globular protein structure. Furthermore, the number of moles of hapten coupled per mole of protein can play an important role in determining the success of antibody production as well as the particular site of coupling in reference to the drug molecule. The resulting conjugate can then be used to produce drug-specific antibodies.

A second consideration which must occur early in the development of an RIA procedure is that of availability of an appropriately labelled drug tracer. This may require custom synthesis since relatively high specific activities are desirable. Associated with the tracer are considerations of stability, purity and total cost.

With the availability of a suitable antigen, one can proceed to raise antisera using standardized immunization schedules with whole animals such as sheep or rabbits. The resultant antisera obtained will be polyclonal in nature since the antisera will have arisen from a whole host of antibody producing cells. Alternately, one can make use of tissue culture

techniques and cell hybridization so as to raise monoclonal antibodies. The utility of both these techniques will be considered further in specific work concerning the development of RIA procedures for 7-hydroxyCPZ.

Finally, once a drug-sensitive antiserum has been produced, it must be characterized both in terms of its sensitivity and specificity. In many cases, specificity can only be delineated with respect to those drugs or metabolites which are available. A final assessment which should ideally always be carried out is confirmation of the quantitative aspects of the RIA procedure by comparison to a chemical method of proven specificity.

Our Approach to Drug Specific RIA

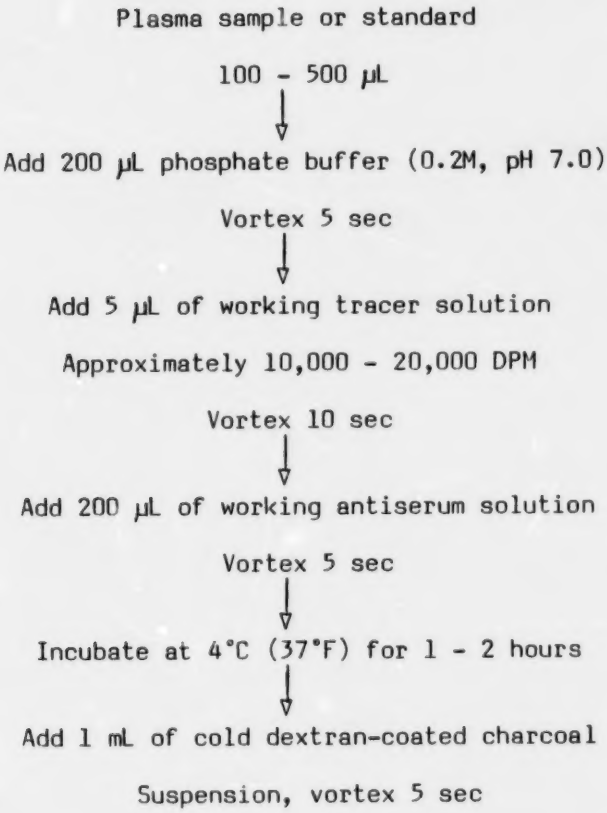
In order to successfully develop RIA procedures of some antipsychotic drugs, we have developed an approach which attempts to always ensure some level of success. In this approach, the first important step is to critically examine the structure of the drug in question considering potentially interfering metabolites in terms of not only their presence but also in terms of their known relative concentration in plasma. Subsequently, the required hapten is synthesized and characterized by NMR, IR and mass spectrometry. The pure hapten is then coupled to bovine serum albumin, or in more recent experiments thyroglobulin, to yield conjugates which have both low (30 moles/mole protein) and high (>30 moles/mole protein) moles coupled. Again, the integrity and number of moles coupled are critically examined using a number of chemical procedures.

In the development of polyclonal antibodies, at least four rabbits are immunized with the respective conjugates on a biweekly basis using subcutaneous injections in adjuvant. Once the immunization schedule commences the titre of the various rabbit sera is checked routinely. Only those rabbits producing a titre of greater than 1:250 are considered useful. Prior to the collection of antisera, at least one intravenous booster injection is given and the serum collected one week later.

The harvested antiserum is then evaluated in order to determine optimum RIA conditions in terms of temperature and time of incubation.

Typically, the procedure shown in Figure 1 is followed. Once optimal conditions are reached, the cross reactivity of the antiserum is checked against available metabolites and structurally similar drugs. In those cases where significant cross reaction with metabolites occurs, the possibility of a selective extraction procedure prior to RIA is considered. A case in point is in a developed RIA procedure for fluphenazine (Dudley et al., 1983).

Figure 1


Plasma sample or standard
100 - 500 µL
Add 200 µL phosphate buffer (0.2M, pH 7.0)
Vortex 5 sec
Add 5 µL of working tracer solution
Approximately 10,000 - 20,000 DPM
Vortex 10 sec
Add 200 µL of working antiserum solution
Vortex 5 sec
Incubate at 4°C (37°F) for 1 - 2 hours
Add 1 mL of cold dextran-coated charcoal
Suspension, vortex 5 sec

Incubate at 4°C for 10 min, centrifuge (4°C, 10 min, 3000 rpm), decant super-natant into scintillation vial containing 10 mL of scintillation fluid, mix and count.

Using this technique, it was found that after oral administration of fluphenazine, the measured plasma levels of fluphenazine were two to five times higher than the levels measured using a sensitive and specific GC/MS assay for fluphenazine (McKay et al, 1983). The apparent problem was that the RIA technique was also measuring a significant level of some cross reacting metabolite(s) of fluphenazine. This problem could be alleviated by extracting the plasma samples for fluphenazine prior to RIA analysis and thereby selectively leaving behind the cross reacting metabolite(s).

Examples of Drug-Specific RIA

In many cases, the development of drug-specific RIA procedures was necessitated due to the need for highly sensitive methods for the determination of selected drugs or metabolites (especially phenothiazines) and/or the general lack of specific chemical methods for drug analyses suitable for routine clinical therapeutic monitoring.

In these laboratories, the use of RIA techniques dates back to our pioneering work on the analysis of CPZ in plasma (Midha et al., 1979) and has continued to date such that RIA procedures have been developed in our laboratories for the drug and/or metabolites shown in Table I.

TABLE I. Drug-Specific Radioimmunoassays and Related Chemical Assays Developed in These Laboratories

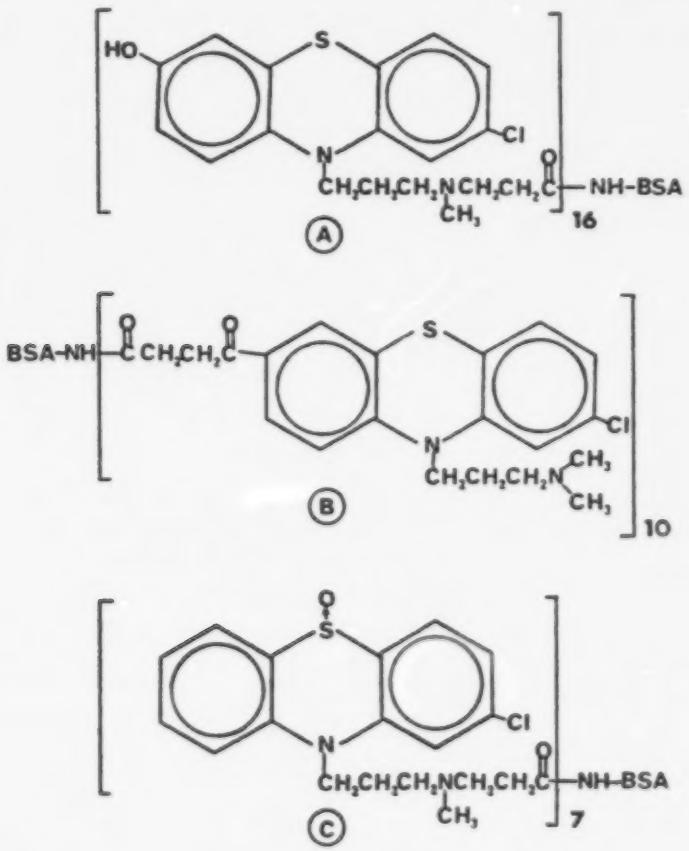
Drug	Approximate Detection Limit (ng/ml plasma) RIA	GC	HPLC
Phenothiazines			
Fluphenazine	0.25	0.08 (MS)	0.50 (EC)
Fluphenazine sulfoxide	0.30	--	--
Prochlorperazine	0.25	0.13 (MS)	0.50 (EC)
Perphenazine	0.25	--	0.10 (EC)
Trimeprazine	0.31	--	0.25 (EC)
Thioridazine	0.40	--	2.5 (UV)
Trifluoperazine	0.15	0.08 (MS)	0.50 (EC)
Trifluoperazine sulfoxide	0.30	--	--
7-Hydroxytrifluoperazine	0.10	--	--
Trifluoperazine N_4-oxide	0.10	--	--
Chlorpromazine	0.15	0.25 (MS)	0.25 (EC)
Chlorpromazine sulfoxide	0.20	--	--
7-Hydroxychlorpromazine	0.50	--	--
Chlorpromazine N-oxide	0.25	--	--
Tricyclic Antidepressants			
Doxepin plus	0.25		2.5 (UV)
N-desmethyldoxepin			2.5 (UV)
Trans-Doxepin	0.15	--	5.0 (UV)
Imipramine/Desipramine	0.10	2.0 (NPD)	--
Amitriptyline	1.0	--	--
Miscellaneous			
d-Ephedrine	2.5	5.0 (ECD)	
l-Ephedrine	2.5		
dl-Fenfluramine	0.30	5.0 (ECD)	--
Chlorpheniramine	0.15	--	2.5 (UV)

A. Chlorpromazine RIA

In reference to CPZ we have now developed immunoassay techniques to monitor CPZ (Midha et al., 1979), 7-hydroxyCPZ (Yeung et al., 1985) and CPZ sulfoxide (Yeung et al., 1983) in patient and volunteer plasma samples. The methods demonstrate adequate sensitivity to monitor the plasma levels after a low single oral dose (50 mg) for up to 24 hours.

In developing the RIA procedures for CPZ and its metabolites, it was necessary to raise antisera to the various protein conjugates shown in Figure 2. The respective conjugates were injected into New Zealand white rabbits and antisera harvested. The cross reaction profiles of the various antisera are shown in Table II. A more specific antiserum for 7-hydroxyCPZ was obtained through the use of mouse monoclonal antibodies (Yeung et al., 1985). The monoclonal antibodies were prepared through the use of the same protein conjugate used in the polyclonal experiments. However, as shown in Table II, the cross reactivities of the antiserum were far superior and was used in subsequent analytical experiments.

FIGURE 2



Protein conjugates for the development of RIA procedures for A, 7-hydroxy CPZ; B, CPZ and C, CPZ sulfoxide.

TABLE II. Cross Reaction* Profiles of Antisera for CPZ and Metabolites

	Antisera			
	CPZ	CPZSO	7-OHCPZ monoclonal	7-OHCPZ polyclonal
CPZ	100%	<1%	<1%	17%
7-OHCPZ	<1%	<1%	100%	100%
CPZSO	<1%	100%	<1%	<1%
CPZNO	<1%	<1%	<1%	<1%
NOR CPZ	100%	<1%	<1%	<5%
NOR CPZ	7%	<10%	ND	ND
7-OH-NOR CPZ	ND	ND	37%	80%
8-OHCPZ	ND	ND	<1%	<8%
6-OHCPZ	<1%	ND	3%	3%

* % cross reaction was calculated by the method of Abraham (J. Clin. Endocrinol., 29, 866, 1969).

Calibration curves for the RIA procedures, based on these antisera, are shown in figure 3 which shows they have sensitivity to at least 0.25 ng/ml plasma. Application of these procedures to the analysis of plasma obtained from a volunteer who received a single 50 mg oral dose of CPZ is shown in figure 4.

FIGURE 3

RIA calibration curves for the plasma determination of CPZ, CPZ sulfoxide and 7-hydroxy CPZ.

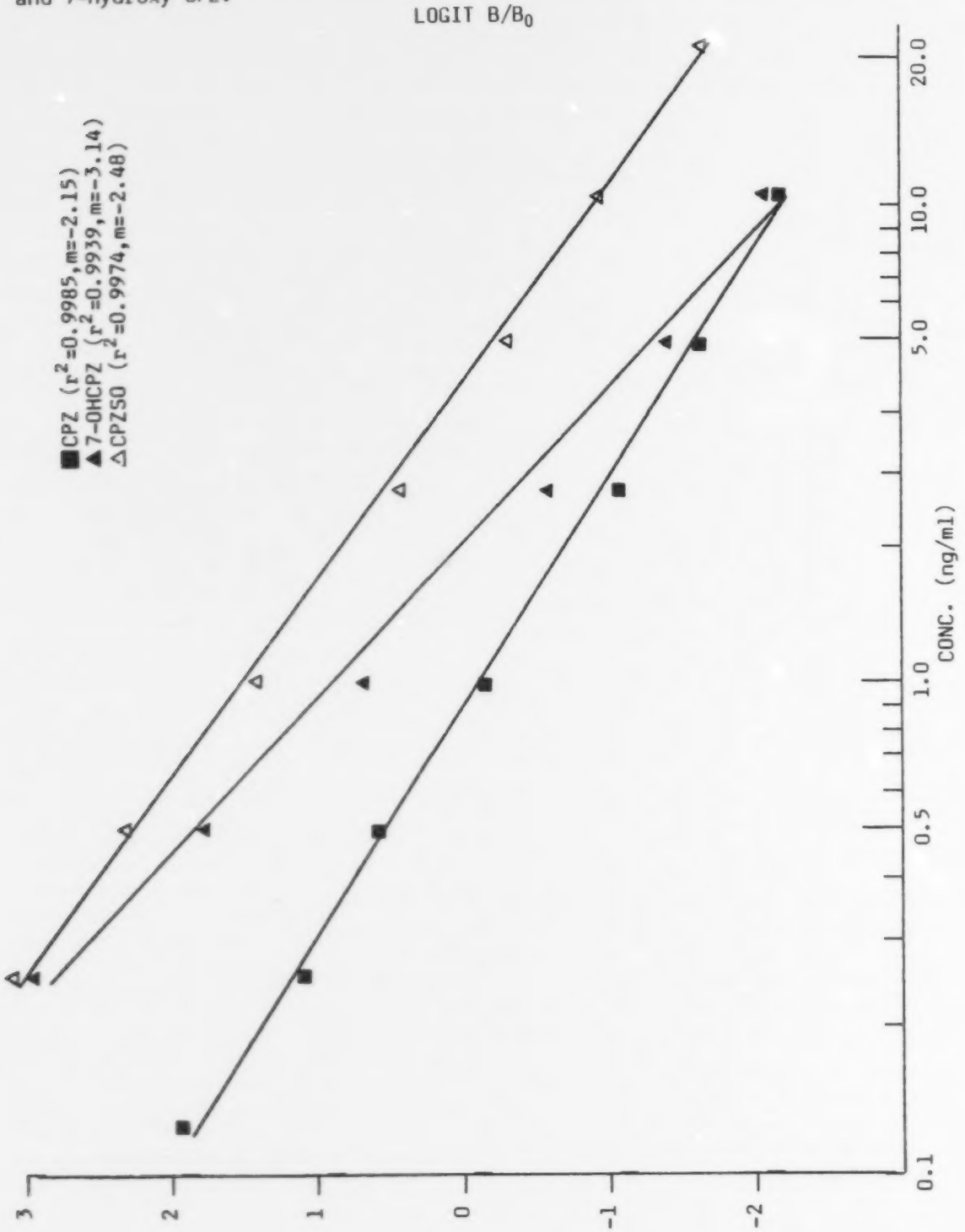
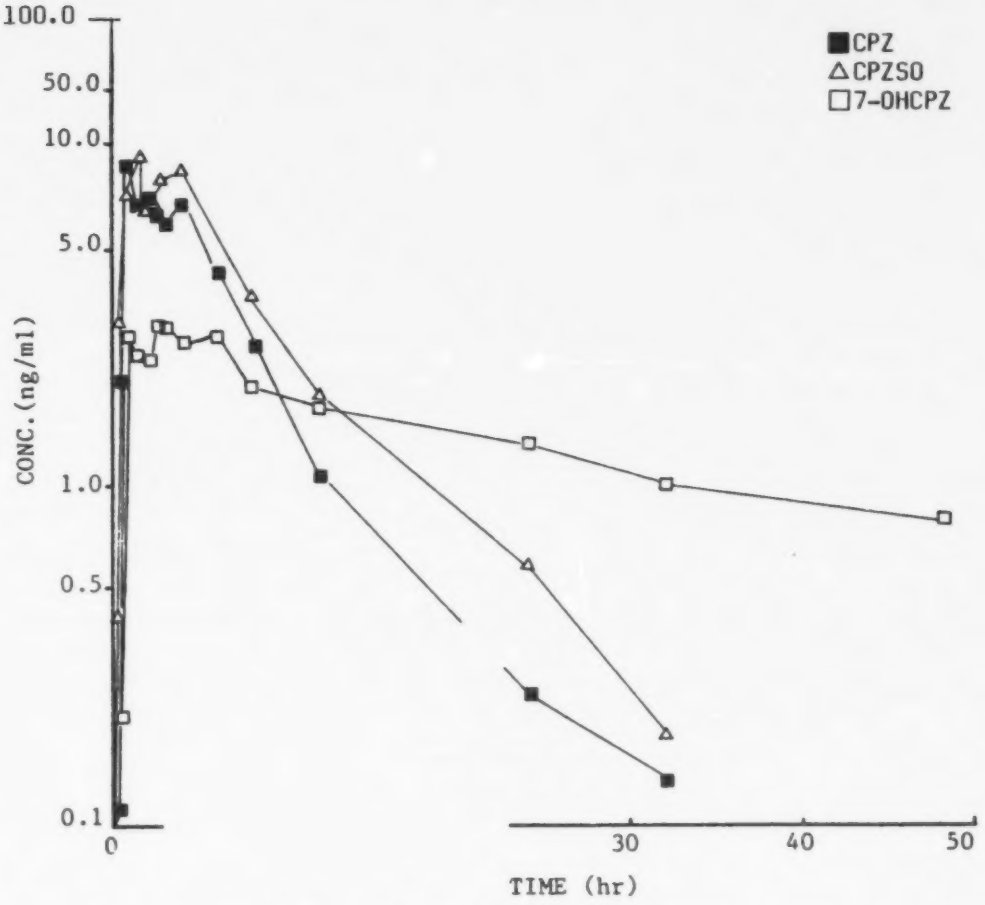


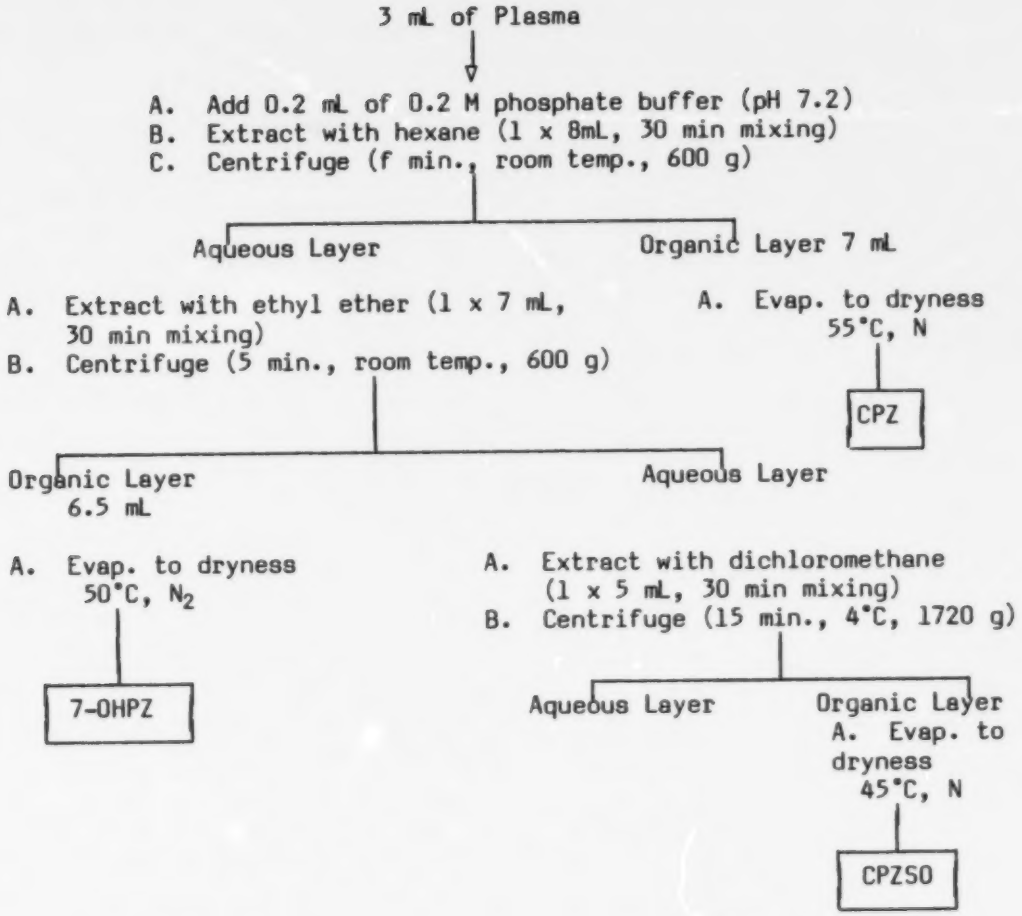
FIGURE 4

Plasma concentration/time profiles for a healthy male volunteer who ingested a single 50 mg oral dose of CPZ.



In order to obtain the data shown in Figure 4, it was necessary to devise a differential extraction scheme for CPZ and these metabolites prior to RIA analysis. The need for such an extraction scheme became evident when the cross reactivities of the various antisera were evaluated by a method other than that of Abraham's (Abraham, 1979). It was found that there was significant interference in the direct RIA procedures if each of the antisera is assessed in the presence of the other metabolites, a situation more closely approximating the conditions in vivo. However, if one differentially extracts CPZ and the various metabolites according to the scheme shown in Figure 5, these interferences are greatly reduced and the data obtained more closely approximate the known or chemically assayed concentrations of the various analytes. This indicates the need for very critical evaluation of biological techniques such as RIA and comparisons of RIA procedures with chemical methods of analysis.

FIGURE 5


3 mL of Plasma
A. Add 0.2 mL of 0.2 M phosphate buffer (pH 7.2)
B. Extract with hexane (1 x 8mL, 30 min mixing)
C. Centrifuge (f min., room temp., 600 g)
Aqueous Layer
Organic Layer 7 mL
A. Extract with ethyl ether (1 x 7 mL, 30 min mixing)
B. Centrifuge (5 min., room temp., 600 g)
A. Evap. to dryness 55°C, N
CPZ
Organic Layer 6.5 mL
Aqueous Layer
A. Evap. to dryness 50°C, N2
7-OHPZ
A. Extract with dichloromethane (1 x 5 mL, 30 min mixing)
B. Centrifuge (15 min., 4°C, 1720 g)
Aqueous Layer
Organic Layer
A. Evap. to dryness 45°C, N
CPZSO

A differential extraction procedure for CPZ, 7-hydroxyCPZ* and CPZ sulfoxide in plasma.

*For analysis of total (free + conjugates), 7-OHCPZ concentrations, 2 mL of plasma was incubated at 37°C with 0.2 mL 1 M sodium acetate buffer (pH 5), 0.1 mL 10% sodium bisulfite solution, and 0.05 mL glusulase preparation (Sigma) for 2 hours, 0.5mL 1 M K_2HPO_4 solution (pH 9.5) was added (final pH 7), 7-OHCPZ was isolated according to the above differential extraction.

Trifluoperazine RIA

The development of RIA procedures for the piperazine type phenothiazine TFP (Midha et al, 1981) have paralleled our earlier developments for CPZ. Thus, antisera have been raised in rabbits which demonstrate sensitivity towards TFP and its major metabolites.

In order to raise sensitive and specific antisera, it was necessary to synthesize the conjugates shown in Figure 6. The cross reactivity profiles of the various antisera are shown in Table III. Here it is evident that the antisera raised for TFP sulfoxide may also be useful in measuring fluphenazine S-oxide levels (Aravagiri et al, 1984). This sort of result can often be observed in RIA and shows the positive reward of checking cross reactivity with as many related compounds as possible.

FIGURE 6

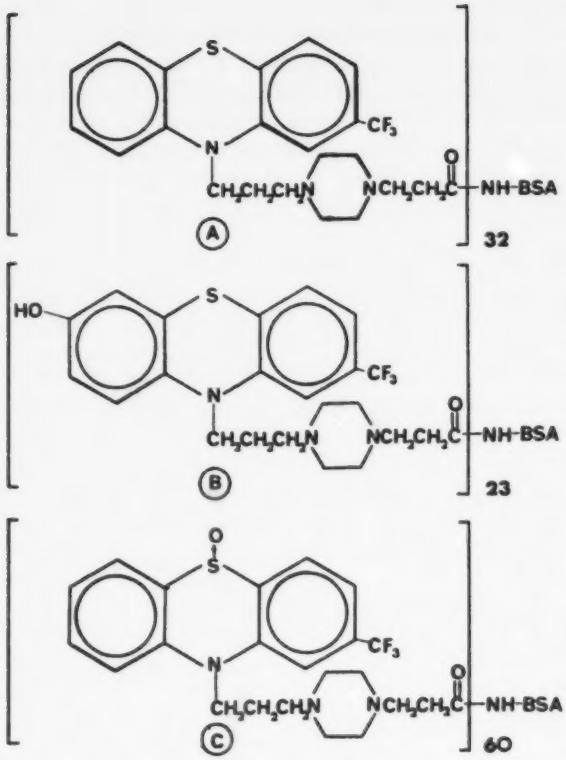


Figure 6 Protein conjugates for the development of RIA procedures for A, TFP; B, 7-hydroxyTFP and C, TFP sulfoxide.

TABLE III. Cross Reaction* Profiles of Selected FP Antisera

	Antisera		
	TFP	TFPSO	7-OHTFP
TFP	100%	<1%	<1%
7-OHTFP	25%	<1%	<100%
TFPN O	ND	<1%	<1%
TFPSO	<1%	<100%	<1%
NORTFP	<1%	<1%	<1%
NORTFPSO	ND	55%	<1%
FLUSO	<1%	62%	<1%

* % Cross reaction was calculated by the method of Abraham (J. Clin. endocrinol., 29, 866, 1969).

The developed RIA procedures are sensitive to at least 0.15 ng/ml with linear calibration curves over the range of 0.15 to 10 ng/ml. Application of these procedures to the analyses of plasma samples taken from a volunteer who ingested an oral 5 mg tablet of TFP demonstrate plasma concentration versus time profiles in which TFP, trifluoperazine sulfoxide and 7-hydroxyTFP could be measured as late as 48 hours after this single oral dose (Figure 7). Further application of these techniques to patient samples being treated with this agent will assist in establishing the relationship, if any, between plasma concentrations of the drug and metabolites, therapeutic response and/or adverse drug reactions.

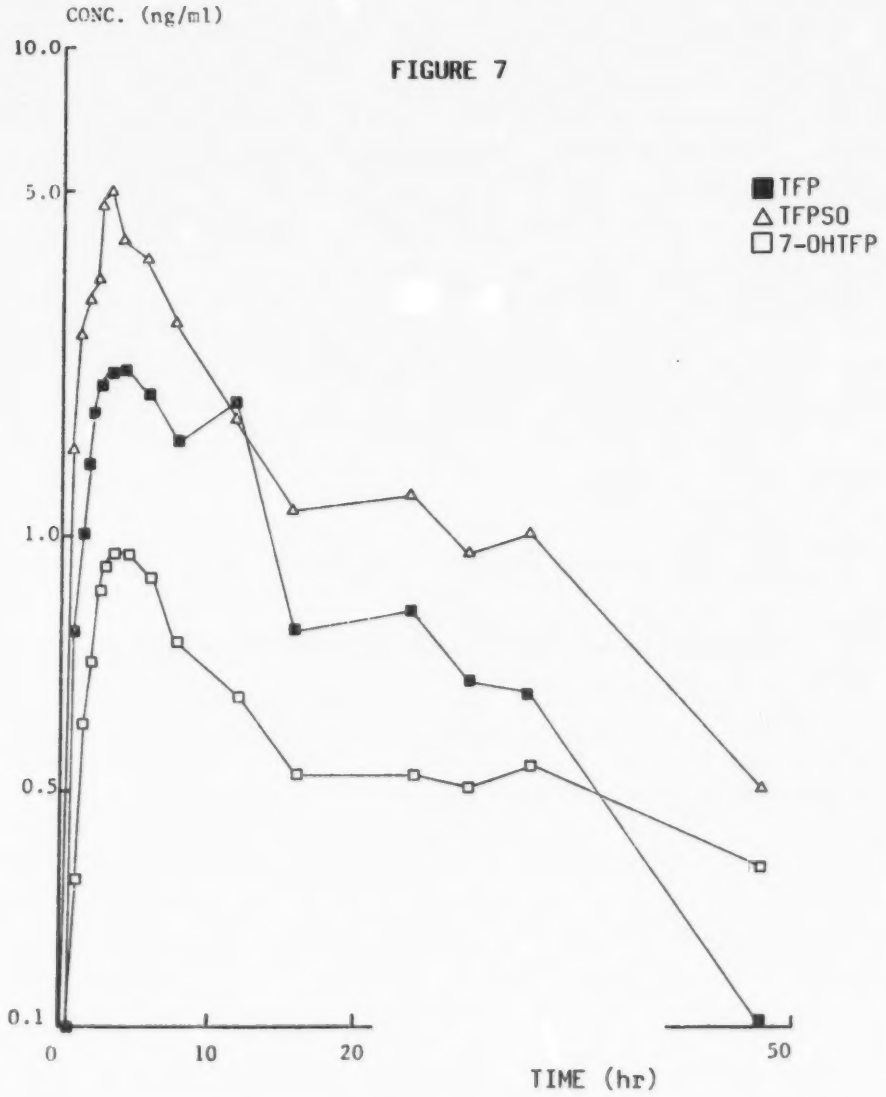


FIGURE 7

Figure 7 Plasma concentration/time profiles for a healthy male volunteer who ingested a single 5 mg oral dose of TFP.

An ultimate test of specificity in any analytical method is in the quantitation of geometric or stereo isomers. In this regard, RIA procedures incorporate the often noted specificity of biological systems. An example of this specificity is in an RIA procedure developed in our laboratories for doxepin. Antisera was raised in rabbits to a conjugate of trans-N-desmethyl doxepin acrylate and BSA. The resultant antiserum shows only 9% cross reactivity with cis-doxepin and has adequate sensitivity to measure doxepin plus N-desmethyl doxepin for up to 24 hours after a single oral dose of 25 mg.

Future Trends in the Use of RIA

With the advancements in molecular biology and the widely increased use of cell fusion techniques the use of monoclonal antibodies for the development of RIA procedures has become quite reasonable. However, this technique is still not without its own set of difficulties since up until now, the use of monoclonal antibodies has most often involved insoluble antigens. In the case of drug-specific antibodies, the antigens are often soluble, therefore, making them to a degree less immunogenic and consequently dropping the number of position clones which can be screened. In our experience, we have had to deviate somewhat from conventional monoclonal antibody techniques in order to get adequate numbers of positive clones. In this regard, we are only at preliminary stages of the use of this technique but see already the potential it holds in developing highly specific antisera to not only important psychoactive drugs but also many of their important metabolites. The use of RIA in studies involving stereoselectivity has only begun to be exploited.

With such techniques, it will be possible to investigate stereoselective metabolism of drugs and evaluate the importance of such metabolites to therapeutic response.

REFERENCES

Abraham, G.E., Solid-phase radioimmunoassay of estradiol-17-β, J. Clin. Endocrinol. Metab., 29 866, 1969.

Aravagiri, M., et al. Radioimmunoassay for the sulfoxide metabolite of trifluoperazine and its application to a kinetic study in humans, J. Pharm. Sci., 73, 1383, 1984.

Dudley, J. et al. Correlation of fluphenazine plasma levels versus clinical response in patients: a pilot study, Prog. Neuro-Psychopharmacol. Biol. Psychiat., 7, 791, 1983.

Hawes, E.M., et al. Therapeutic monitoring of chlorpromazine III: Minimal interconversion between chlorpromazine and metabolites in human whole blood under physiological conditions, Ther. Drug Monitoring, (submitted), 1985.

Hubbard, J.W. et al. Therapeutic monitoring of chlorpromazine I: Pitfalls in plasma analysis, Ther. Drug Monitoring, (in press), 1985.

MacKay, A.V.P. et al. The relationship of plasma chlorpromazine to its 7-hydroxy and sulfoxide metabolites in a large population of chronic schizophrenics, Br. J. Clin. Pharmac., 1, 425, 1974.

McKay, G., et al. Subnanogram determination of fluphenazine in human plasma by gas chromatography-Mass Spectrometry, Biomed. Mass Spect., 10, 550, 1983.

Midha, K.K. et al. Kinetics of oral trifluoperazine disposition in man, Br. J. Clin. Pharmac., 15, 380, 1983.

Midha, K.K. et al. Kinetics of oral fluphenazine disposition in humans by GC-MS, Eur. J. Clin. Pharmacol., 5, 709, 1983a.

Midha, K.K. et al. Stereospecific Radioimmunoassays for l-Ephedrine and d-Ephedrine in human plasma, J. Pharm. Sci., 72, 736, 1983b.

Midha, K.K., et al. Radioimmunoassay for trifluoperazine in human plasma, Br. J. Clin. Pharmac., 12, 189, 1981.

Midha, K.K., et al. Radioimmunoassay for chlorpromazine in plasma, Clin. Chem., 25, 166, 1979.

Oliver, G.C. et al. The measurement of digitoxin in human serum by radioimmunoassay, J. Clin. Invest., 47, 1035, 1968.

Yeung, P.K.F., et al. A comparison of two radioimmunoassays for 7-hydroxychlorpromazine: rabbit polyclonal antibodies versus mouse monoclonal antibodies, J. Pharmacol. Exp. Ther., (in press), 1985.

Yeung, P.K.F., et al. A study of the kinetics of chlorpromazine sulfoxide by a specific radioimmunoassay after a single oral dose of chlorpromazine in healthy volunteers, J. Pharm. Exper. Ther., 226, 833, 1983.

ELECTRO-PHYSIOLOGICAL DIAGNOSTIC TECHNIQUES

THE CONVERGENCE OF NEUROPSYCHOLOGICAL AND EEG MEASURES OF CEREBRAL DISORGANIZATION IN THE ENDOGENOUS PSYCHOSES

PIERRE FLOR-HENRY, M.D.
University of Alberta

ABSTRACT

Neuropsychological studies of the endogenous psychoses indicate the presence of progressively increasing cerebral dysfunction: least in depression, intermediate in mania and maximal in schizophrenia. Remarkably, the same progression is seen with certain neurophysiological indicators: EEG coherence and power (8-13 Hz). In addition, fundamental EEG factors derived by principal component analysis from a large sample of normal volunteers will be described (1. - Power, 2 - Coherence, 3 - Complex phase relationships, 4 - Log. right/left power ratios, 5 - Oscillations) and contrasted with the abnormal neuro-physiological configurations seen in a group of 19 exhibitionists.

Neuropsychological

Consecutive patients who satisfied the Taylor and Abrams criteria for schizophrenia, mania and depressive psychoses were included in the study, as well as neurotic depressives (Taylor and Abrams, 1975; Carney et al., 1965). The subjects were all unmedicated and dextral. A battery of 15 neuropsychological tests generating 29 neuropsychological variables and the Wechsler Adult Intelligence Scale (WAIS) were administered. The tests were: Speech Sound Perception, Trail Making Test A and B, Memory for design, Coloured Progressive Matrices (Raven), Category Test, Finger Localization, Finger Tapping Test, Seashore Rhythm, Oral Word Fluency, Purdue Pegboard Test, Face-Hand Test, Dynamometer, L.J. Tactile Recognition, Written Word Fluency, and Wisconsin Card Sorting. For comparative and validation purposes analogous data from 134 age and sex matched normal controls were available. The demographic data for the psychiatric groups were as follows:

	Schizophrenia	Mania	Depression (psychotic)	Depression (neurotic)
Age	31.8	39.5	37.2	34.3
Sex (% male)	51%	50%	47%	43%
Education (years)	10.5	12.3	13.5	12.8
WAIS FSIQ	88.6	99.2	108.3	110.4

Ratings of the neuropsychological test profiles were made according to the methods derived by (Reitan and Davison 1974, Selz and Reitan, 1979) and without knowledge of diagnosis. As the next table shows the neurotic depressions were least and the schizophrenias most abnormal:

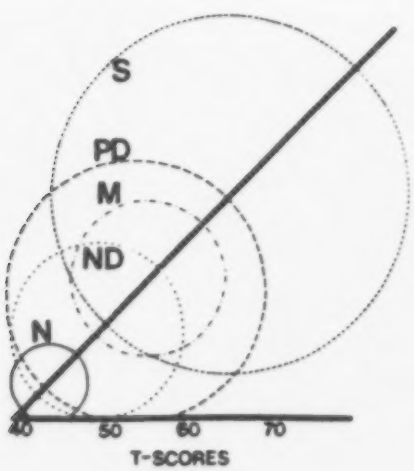
Table 1.

	N	Abnormal (%)
Schizophrenia	32	91
Mania	28	79
Depression (psychotic)	30	57
Depression (neurotic)	24	42
TOTAL	114	68

In order to reduce the number of variables to be considered, a principal component factor analysis (Muliak, 1973) was performed on the 29 variables derived from the 15 neuropsychological tests across the four patient groups. Seven factors accounting for 73.1% of the total variance were rotated to orthogonal simple structure according to the varimax criterion (Kaiser, 1958). Before rotation the first, most important single factor, was a factor relating to severity of neuropsychological dysfunction. Figure 1 shows the average score of each diagnostic group on this factor: standardized (T-scores) with a centre of 40, a mean of 50 and a standard deviation of 10.

FIGURE 1

Group Severity



There is a clear increasing continuum of cerebral dysfunction: least in neurotic depression, then with increasing severity in psychotic depression, mania, maximal in schizophrenia. The enormous variability within the schizophrenia group is noteworthy, as is the considerable degree of overlap with psychotic depression, and the much more homogenous nature of mania which lies intermediate between psychotic depression and schizophrenia (Flor-Henry, Fromm-Auch and Schopflocher, 1983).

EEG Measures

In a multivariate analysis of the EEG in 63 psychotic depressives, 75 manic and 53 schizophrenic patients who satisfied both the research criteria of Feighner et al. (1972) and Taylor and Abrams (1975), coherence emerged as an important factor. All the patients were unmedicated and dextral and were compared to each other and to 60 normal controls. The analysis is reported for temporal (T_3, T_4) and parietal (P_3, P_4) derivations in a unipolar montage against Cz mid-line reference. The subjects were studied during rest (eyes open and closed), during two verbal cognitive tasks (Vocabulary subtest of the WAIS, Oral Word Fluency) and during spatial processing (Block Design subtest of the WAIS). Only the alpha frequency band (8-13 Hz) is considered in the statistical analysis. A principal component analysis of the 54 derived EEG measures for every subject and all conditions in the four groups produced four factors which accounted for 41% of the total variance. Orthogonal quartimax rotation showed that they were: 1- Coherence; 2- Power and Variance; 3- Phase (antero-posterior); 4- A complex factor relating right/left temporal power ratio to coefficient of variation and to both back-front coherence and phase relationships.

Figure 2 indicates the (standardized) group means on the coherence factor. We see that the normals have the highest coherence, followed by the psychotic depressions, with mania and schizophrenia exhibiting the smallest intra and inter-hemispheric coherence (Flor-Henry, Koles and Sussman, 1983).

FIGURE 2

Multivariate EEG Analysis of the Endogenous Psychoses

Means on general coherence by diagnosis

In another analysis of the EEG in these groups we considered, in addition to the power, variance, coherence and phase variables already treated, a new measure which we introduced to estimate the rate of right/left hemispheric energy fluctuations through time (Flor-Henry and Koles, 1981). Thus the number of right/left hemispheric energy shifts per second, or oscillations, are calculated 15 times per second for two to three minute epochs. A principal component analysis of the schizophrenias (n=53), manias (n=75), psychotic depressions (n=53) and normal controls (n=60) in the five conditions (EO, EC, Voc., OWF and BD) produced seven factors which accounted for 70% of the total variance. After varimax orthogonal rotation they were: 1- Power and Variance; 2- Right/left hemispheric energy oscillations; 3- Coherence; 4- Phase (left→→→ right) hemisphere; anterior→→→ posterior: (right→→→ left) hemisphere; posterior→→→ anterior; 5- Log. right/left temporal power ratio; 6- Phase: left temporal→→→ parietal, right and left: right temporal→→→ parietal, right and left; 7- Log. right/left parietal power ratio. These factors were transformed into seven corresponding scales according to the allocation of binary weights on the basis of high or low loadings on the factors, the weights of measures being adjusted to have equal means and variance (0 mean, unit variance). A discriminant function analysis showed that the various groups could be differentiated from each other: the critical factor depending on the condition. In the eyes closed condition the schizophrenics had significantly less power than the normals. As Table 2 indicates we see the following reduction in alpha power: maximal in normals, gradually diminishing in depression and mania, minimal in schizophrenia (Flor-Henry, Reddon and Koles, in preparation, 1984).

Conclusion

The findings in these neuropsychological and statistical quantitative EEG investigations of normals, depression, mania and schizophrenia are remarkably consistent and convergent. With quite different measures of cerebral organization: neuropsychological parameters, EEG coherence and EEG power it is shown that the severity of cerebral dysfunction is least in depression (who are closest to normals), intermediate in mania and maximal in schizophrenia. Mania and schizophrenia, on these indices of brain disorganization are very similar to each other, with a lesser intensity of dysfunction in mania.

TABLE 2

EEG Analysis of Psychosis (8-13 Hz)

(Schizophrenia, N=53, Mania, N=75, Depression, N=63, Normals, N=60)

Discriminant Function Analysis

Task			Major Discriminatory Variable
Eyes Closed Overall $\bar{X}$: +2.85	*Schizophrenics	+0.90	Power and Variance
	Manics	+1.97	
	Depression	+3.26	
	*Normals	+5.25	

Schizophrenia: EEG characteristics in males and females (P_3,P_4,T_3,T_4, -vs C_z - EO, EC, VOC, OWF, BD)

REFERENCES

Carney, M.W.P., Roth, M. and Garside, R.F. (1965) The diagnosis of depressive syndromes and the prediction of E.C.T. response. British Journal of Psychiatry, 111, 659-674.

Feighner, J.P., Robins, E., Guze, S.B., Woodruff, R.A. (Jr.), Winokur, G., Munoz, R. (1972) Diagnostic criteria for use in psychiatric research. Archives of General Psychiatry, 26, 57-63.

Flor-Henry, P., Koles, Z.J. (1981) Studies in right/left hemispheric energy oscillations in schizophrenia, mania, depression and normals. Advances in Biological Psychiatry, 6, 60-67.

----------, Fromm-Auch, D., Schopflocher, D. (1983) Neuropsychological dimensions in psychopathology. In P. Flor-Henry and J. Gruzelier (eds) Laterality and Psychopathology, Elsevier Science Publishers, North Holland, 59-82.

----------, Koles, Z.J., Sussman, P.S. (1983) Multivariate EEG analysis of the endogenous psychoses. Advances in Biological Psychiatry, 13, 196-210.

Kaiser, H.F. (1958) The varimax criterion for analytic rotation in factor analysis. Psychometrika, 23, 187-200.

Muliak, S.A. (1973) The Foundations of Factor Analysis. McGraw-Hill Book Company, New York.

Reitan, R.M. and Davidson, L. (Eds) (1974) Clinical Neuropsychology: Current Status and Applications. New York: Wiley.

Selz, M. and Reitan, R. (1979) Rule for neuropsychological diagnosis: classification of brain function in older children. J. Con. Clin. Psychol., 47, 258-264.

Taylor, M.T. and Abrams, R. (1975) Manic-depressive illness and good prognosis schizophrenia. American Journal of Pshychiatry, 132, 741-742.

MOLECULAR BIOLOGY, CELL BIOLOGY, DRUG RESEARCH AND HUMAN BEHAVIOUR

MOLECULAR BIOLOGY, CELL BIOLOGY, DRUG RESEARCH AND HUMAN BEHAVIOUR

J. Fraser Mustard, M.D.
The Canadian Institute for
Advanced Research

ABSTRACT

Examples of the rapid pace and complexity within today's research in the biological sciences are described. The need for change and adjustment within the institutional structures involved in this research is then illustrated. Several problems and challenges are defined and discussed, particularly those of funding, the difficulties of cross-disciplinary interaction, the lack of a mechanism for inter-university co-operation. Possible solutions and options are presented.

The pace of development of new knowledge and its application in society has increased substantially in the latter part of this century. Already the impact of these developments is recognized in a number of areas such as the emerging technological application of "artificial intelligence" in the work place. Those involved in this change are well aware of the effect on the nature of work and productivity, and the new approaches to research and development that need to be taken to make use of new opportunities.[1]

Research and development in the human biological sciences have also moved at an extraordinarily rapid pace in recent years. One has only to consider the successful development of a cell line to synthesize the missing protein in hemophilia to grasp the pace and sophistication of the research in the biological sciences. There is an awareness in some quarters of the concentration of resources that are necessary to keep competitive research programs functioning in these rapidly moving fields.[2] The final stages of the isolation of the factor VIII protein (the missing protein in hemophilia), characterization of its key amino acid sequences, the creation of DNA copies and cloning of cells to produce factor VIII, took place in about three years and involved more than 40 Ph.D. scientists. We shall obviously see further extensive application of molecular and cellular biology to disorders associated with the blood, partly because of the easy access to this body system and the extensive knowledge base which has already been developed in this area.

When one turns to the neurological system and behaviour, the problem is more complex because, as all of you are aware, we do not yet have the detailed knowledge base, in terms of structure and function, which has been developed for hematology. However, the emergence of new tools for studying the human brain has begun to give us the opportunity to extend our understanding of this far more complex body system. The application of the techniques of molecular and cellular biology to the function of nervous

tissue in relation to behaviour, intelligence and memory, provides a major research opportunity for the future, and the results from this research will create the need to make some difficult decisions about the use of the new knowledge that will be acquired.

I would like to cite a practical example of the developments in cell biology that relate to the nervous system and behaviour. Inositol phospholipids represent a fairly unique, minor phospholipid group in the membrane of cells. Interest in these phospholipids and cellular function initially developed in relation to nerve transmission. We now know from studies on blood cells and some other cells that the inositol phospholipids are probably the critical membrane phospholipids in the response of most cells to stimuli[3] and are therefore closely involved in cell-to-cell communication, which is of course an important aspect of the function of neurological tissue. Lithium, which was applied to the treatment of manic depressive illness, is now known to affect inositol phospholipid metabolism. Lithium inhibits the degradation of the inositol phosphates[3,1] which are the breakdown products of inositol phospholipids. The inositol triphosphate that is formed by the action of phospholipase C on triphosphatidyl inositol plays a key role in the mobilization of intracellular calcium and the activation of proteins that carry out processes within cells that are involved in cell secretion and movement. We also know that the other product of the action of phospholipase C is 1,2-diacylglycerol. This compound activates a protein kinase that is also important in the biological responses of cells and this product acts synergistically with inositol triphosphate. Is this action of lithium on the metabolism of phosphatidyl inositol a clue to understanding how lithium exerts its effects? If so, how can we relate further knowledge of this metabolic pathway and its modulation to our understanding of disorders of behaviour and their treatment?

Finally, we know that when these inositol phospholipids are stimulated, they can release arachidonic acid which is transformed into a variety of prostaglandins and the thromboxanes that also modulate cellular function.[3] The fatty acid composition of these phospholipids is influenced by diet. In the cardiovascular system we know that diets rich in polyunsaturated fatty acids such as eicosapentaenoic acid dramatically change the responses of these phospholipids and influence cellular functions that are involved in the development of vascular disease. A diet that includes large amounts of eicosapentaenoic acid is believed to be one of the reasons that the Eskimos of Greenland show such a low incidence of cardiovascular disease. There has not yet been much transfer of this knowledge to the study of other body systems such as the nervous system. It is clear, however, that because of the general application of new knowledge in cell biology, there will be application of this knowledge concerning the inositol phospholipids and arachidonic acid metabolism to the nervous system. The disciplines involved in the development and application of this knowledge to the nervous system will include bio-chemistry, cell biology, neuroscience, psychiatry, psychology, nutrition and pharmacology. How do we create the manpower concentration to produce the key intellectual dynamic that is necessary to facilitate the cross-disciplinary interaction that will be required for effective research in an area such as this?

I have presented this story because it demonstrates the extraordinary excitement of modern research in the biological sciences, but it also illustrates its pace and complexity. Today the traditional disciplines that were developed before we had our current knowledge base are probably less and less useful in setting up research approaches for the kinds of problems that can now be studied. Thus, there is a need for change and adjustment in the institutional structures to meet the new opportunities. We are seeing the emergence of new approaches to the study of problems in a number of areas including the neurosciences. Perhaps a good example of such adaptation is the MacArthur Foundation's support of a program[4] to bring together several different disciplines in a number of institutions to tackle the complex research problems in the behavioural sciences. This is a science network. A similar development is the initiative of the Rockefeller University to form a network in the neurosciences on the eastern seaboard of the United States. These complex fields require linkages in research resources which have previously not been developed, in order to bridge the gaps between disciplines, create manpower concentrations and diminish the existing barriers to the intellectual activity that is necessary for progress in these fields of research.

A characteristic of modern research is the manpower and resource concentrations that are necessary for it to be effective. All of you are familiar with the demonstrated value of positron-emission tomography and nuclear magnetic resonance in the study of the nervous system. These tools, and the more sophisticated next-generation instruments, are powerful aids in our study of the functions of the brain. These facilities are very expensive and therefore, will have to be in centralized locations where they can be used by many scientists. How will we develop the linkages to make these facilities effective?

The disciplinary interaction which is necessary in fast-moving research fields requires the creation of a manpower base that cannot be easily developed in a single institution in Canada. Our universities are primarily funded for a mass education role, with basic research as a subsidiary function. The current research fields are moving at such a pace that in order to achieve the research concentrations to be competitive and successful, new research funding policies should be developed to meet the new challenges. This is now difficult to do from the funding base provided for the mass education function of universities. Can we really sustain the kind of basic and applied research that is needed within our existing institutions when they have very little capacity to meet the overhead costs of the research they undertake? Will we not have to change the structure of the support mechanisms for research to make it possible to move with the changing nature of modern research?

The linkages between basic, applied and developmental research have also been changing. The time frame for the development of new knowledge and its application has become very short in some fields. Are our institutional arrangements for basic, applied and developmental research in

Canada appropriate for the changes that are taking place? Have our policies towards the applied health industry handicapped Canada and prevented effective industrial research in the biological sciences and the development of drugs for use in man? The policies were developed in a previous era. Should we not now change those policies if we want to have a strong linkage between basic and applied research and the development and the effective application of new knowledge in the health field?

The challenge for our country, indeed, for all countries is clear. If we do not move to make our institutional and funding arrangements more relevant to modern research, we will eventually lag far behind in developments. Rather than capitalizing on where we are and moving ahead, we will be doing derivative work on research pioneered in other countries, and struggling to keep up. If this happens, we will be unable to retain the talent in the country that is necessary to apply the knowledge that is emerging and the country may find it difficult to make the adjustments facing all societies as a result of the development of this knowledge.

All of you concerned with the mental health field fully appreciate what failure to keep pace with the development of new knowledge and its application in society means to one of the biggest and most complex health burdens in our society - mental illness.

The Problems and Solutions

A recent editorial in Science[5] addressed the question of new approaches to research. The writer argued:

> It is vital that the universities take a serious look at their futures. New directions are opening while universities are merely plying their daily business. Charting a course of institutional development for the next generation of research universities is perhaps the most urgent need facing higher education at this time. Networks of research and instruction in high technology may be one way to begin, with alliances developing through the interaction of faculty and students ... consortiums may be the outline of what universities will become in the next century.

In the light of the preceding discussion, the problems facing Canadian research may be summarized as follows:

- Universities are funded by provincial legislatures, whose natural tendency is to resist the concentration of resources in a few selected institutions. Without such concentrations, however, it is difficult to build the "critical mass" of research talent necessary to be competitive internationally.

- Under existing funding arrangements, university income is primarily tied to its teaching function. Higher enrollment leads to higher revenue. This makes it difficult to achieve concentrations of research talent, because the manpower needs for research are not directly related to the number of undergraduate students.

- Under the existing policy of the federal granting councils, the universities must absorb the indirect costs of research. Since their own real incomes have been shrinking in recent years, this has led to a marked deterioration in both their capacity to support the indirect costs of research and the quality of research facilities.

- The relationship between the universities and industrial research and development (i.e. between basic and applied research) is not well developed. As a result, the level of non-governmental support for university research is well below that found in other countries.

- Because of their discipline-based "departmental" structures, the universities find it difficult to facilitate the kind of cross-disciplinary interaction that is an essential part of many new fields of research.

- There is no mechanism to facilitate inter-university co-operation, particularly across provincial boundaries. Without such a mechanism, it is difficult to achieve the pooling of resources necessary to achieve concentrations of research talent.

Toward A Solution

The situation would be greatly improved if both federal and provincial governments co-operated in the formulation of new funding arrangements for universities, particularly arrangements that would require them to fund the indirect as well as the direct costs of research.[2,6,7]

The establishment of what has often been referred to as "centres of excellence" within universities could help ease some of the problems.[6] Such university-based research groups not only help to sustain competitive, high-quality research, but also provide a means for developing new highly-qualified manpower in areas of great need and rapid growth. Centres of excellence that are remote from the universities will have less flexibility and will divert resources from the education of the trained manpower needed by society.

Even if all these arrangements were put into place, however, a number of important problems would remain, among them:

- how to achieve concentrations of research resources that cannot easily be achieved within existing institutions,
- how to promote cross-disciplinary interaction,
- how to encourage productive interaction between universities and the private sector,
- how to achieve competitive high-quality research in sectors of new research fields.

There are a number of options which should be considered in trying to address these problems:

1. A New Research University- An attempt could be made to create a new "national" university, located in one geographic site and focused on advanced research in the resource intensive fields and graduate education, rather than undergraduate education. This would permit the concentration of resources and, if properly structured, the kind of cross-disciplinary research required by many of the emerging fields of research. The establishment of such a university, though, would require a massive capital investment as well as a substantial operating budget. In addition, the concentration of intellectual as well as financial resources could weaken existing institutions and attract a great deal of opposition from the universities.

2. Research Networks- Another option would be to form -- at the initiative of the federal granting councils or the universities themselves -- a series of research networks that would bring talented researchers in various parts of the country together to work on problems of common interest. These networks might also share a common research facility, along the lines of TRIUMF. The arrangements that have been most successful have been for research requiring unique physical facilities. Central physical facilities, however, primarily serve a group of diverse scientists and do not usually create a concentration of resources focused on selected sectors of research fields. There are, in addition, the problems universities have in establishing networks because of the various demands placed on them and the fact that their academic structure is not designed for the development and operation of networks.

3. Private Network Institutes- Such institutes would seek to incorporate the benefits of the other options without being constrained by their limitations. They would build upon the existing centres of research strength in the country, rather than trying to replicate them in one place. On the other hand, they would have the independence to select areas -- and people -- on which to focus resources. They could be goal-oriented and responsive to the needs of Canadian society. This option has the potential to deal -- in selected fields -- with new research initiatives,

the problems of concentration of resources, cross-disciplinary interaction and links with the private sector. Thus it should have, in selected sectors, an amplification effect on the existing research base and the capacity to initiate new research approaches which are difficult to introduce in existing institutions. This is particularly true for research fields where innovation and concentration of resources are pivotal for future development.

By identifying opportunities, mobilizing the resources needed to meet them, and making new arrangements, we can stimulate and focus Canadian research efforts and contribute significantly to the new research thrusts that are important for the country's intermediate and long-term economic and social development.

REFERENCES

1. Bell, D., The coming of post-industrial society, Books, Inc,. New York, 1976

2. The Commission on the Future Development of the Universities of Ontario. Ontario Universities: Options and Futures. Ontario Government Bookstore, Toronto, 1984.

3. Berridge, M.J., Inositol triphosphate and diacylglycerol as second messengers. Biochem. J., 220, 345-360, 1984

3.1 Berridge, M.J., Downes, C.P., and Hanley, M.R., Lithium amplified agonist-dependent phosphatidylinositol responses in brain and salivary glands. Biochem. J., 206 587-595, 1982

4. The John D. and Catherine T. MacArthur Foundation. Report on 1980 and 1981. Marquette Building, Chicago.

5. Vandiver, F.E. - Universities: The next iteration? Science, 225: 791, August 24, 1984.

6. Johnson, A.W. - Giving greater point and purpose to the federal financing of post-secondary education and research in Canada. A report prepared for the Secretary of State of Canada, 1985.

7. Editorial, "Boosting University Research", The Toronto Star, March 24, 1985.

RESEARCH AND DEVELOPMENT ISSUES IN CANADA

ANDREW J. MORIARITY, M.D.

Biomedical Resources/Medi-Edit Ltd.

Some of us naively believed that description of the DNA genetic code some 20 years ago would solve all our biomedical problems. However, this was not the case - could not be the case - no more so than the dictionary represents Shakespearian literature.

The ultimate of course in biotechnology is human gene therapy, and we are now closer to this than many realize. Just about ten days ago, the National Institutes of Health published guidelines for human gene therapy - 'points to consider' in the Federal Register. Such treatment is now considered feasible for selected rare but deadly diseases. And for those who would consider the idea of human genetic engineering repugnant, I would remind you of a statement by Robert Sinsheimer some 10 to 15 years ago.

'I know there are those who find this concept and this prospect of genetic engineering repugnant - who fear with reason that we may unleash forces beyond human scale, and who recoil from this responsibility. I would suggest to them that they do not see our present situation as a whole. They are not among the losers in the chromosomal lottery that so firmly channels our human destinies. This response does not come from the 250,000 children born each year in the United States with structural or functional defects, of which an estimated 80% involve a genetic component.

As we enlarge man's freedom, we diminish his constraints and that which he must accept as given. The equality of opportunity is a noble aim, given the current inescapable genetic diversity of man. But what does equality of opportunity mean to the child with an IQ of 50?'

It is expected that one or more specific proposals for human gene therapy experiments will be forthcoming within the next 12 months, and there are two types of gene therapy being considered - somatic cell and germ cell therapy. Only somatic cell therapy is being studied at this time, and this will probably involve the in vitro treatment of a patient's bone marrow cells with a retrovirus vector carrying the gene for a specific deficient enzyme. The first candidates will be homozygous for genetic defects in such enzymes, as for Lesch-Myhan disease (hypoxanthine-guanine-phosphoribosyl transferase), and certain severe immunodeficiency diseases (purine nucleoside phosphorylase and adenosine deaminase).

Just as for pharmaceuticals, this will require good animal data, assessing the risk of producing pathological viruses or malignancy. Such data must include documentation of - 1) effectiveness of the delivery system, 2) whether there is sufficient expression in animal marrow cultures, and 3) safety testing relative to pathogenicity, including carcinogenicity of the retroviral vector system. Germ line gene therapy is not seriously being considered at this time.

Recombinant DNA research is described as having four major and immediate clinical impacts - 1) carrier detection in prenatal diagnosis, 2) provision of a better understanding of the molecular basis for genetic disease, 3) production of biological products, and 4) human gene therapy. Biological products already produced include human insulin, human growth hormones and the interferons. The next products are vaccines, blood components, newer hormones and diagnostics. For vaccines, perhaps the ones of greatest interest include rabies and poliomyelitis. Where most paralytic polio in the world today results from use of the live virus product, it is important that a potent non-living component antigen be developed. Recombinant DNA techniques now allow this. With regard to hepatitis, it is now described that the first vaccine to prevent cancer has been developed. A recent paper in Science describes a second hepatitis antigen - one that reacts with antibodies in patients with hepatocellular carcinoma. And, of course, the dream we all have is for an influenza virus vaccine that will provide a broader spectrum of protection against the various epidemic strains of the virus.

As referred to by Dr. Mustard, Factor VIII is an important new recombinant DNA product now under development, as well as serum albumin, new thrombolytic agents, and immunomodulators such as interleukin 2, which will shortly be clinically studied for the treatment of cancer. There are also numerous new neurohormones, and neurobiology has been described as the next major area for exploitation of biotechnology. In diagnostics, DNA probes and monoclonal antibodies will make it possible to diagnose a wide variety of genetic disorders and, of course, are being explored relative to the diagnosis and treatment of cancer.

And it is always sobering to note that, undoubtedly, biotechnology will have its greatest and most significant impact in the development of agriculture.

I thought of a subject that would perhaps be more appropriate to this psychiatric audience, and would like to tell you a little bit about molecular genetics as expressed in neurons. We all know that the central nervous system is made up of a vast network of neurons, and that these are anatomically and functionally unique. It is now possible to isolate the genes that are expressed in neurons and responsible for behavioural changes. However, the human is far too complicated, and we must turn to such animals as the earthworm for guidance. The worm has 273 neurons, all of which are different, but these relate to specific activities, including behavioural changes. However, the best system so far for examining such changes - the genetic equivalent of the fruit fly - is the marine snail. 'Egg laying' behavioural activities can be studied in the snail with in situ hybridization techniques. Biological scientists are able to analyze neuron specific genetic expression that results in morphological and behavioural changes. Here, the overwhelming feature is the 'beautiful' or exquisite integration of activity based on single gene-peptide relationships. The marine snail's behaviour is based on 'sets' of non-overlapping peptide systems. The responsible genes evolve, in turn, to form new peptides.

We all know that behaviour is based ultimately on two considerations - specificity of the 'wiring' and neurotransmitters. Some neurotransmitters act both locally and distally as neurohormones. These neurohormones can act on different target cells, producing different effects. A single promoter-translation initiator system can control more than one hormone - resulting in the packaging of sets of peptide hormones in single vesicles that can be released at the same time. In this way, coordinated synthesis and coordinated release are provided. Beyond this, ambiguous polyprotein precursors result in differing 'combinations' of peptide activity, by simple alterations in the metabolic pathway. And there are functionally inert amino acids that intervene in these polypeptides, providing the chemical basis for evolving new activities. Different combinations of overlapping peptides result in different behavioural patterns, and we can see how complex behaviours are exquisitely assembled by combining the activities of individual units.

I am reminded by this of the writings of Szent-Gyorgi a few years ago, when he described the biochemical basis for organized life. If the human body was nothing more than a bag of biochemicals and enzymes, it would easily run down and extinguish itself within a few hours. The secret of life is the build-up of a beautiful array of indirect or alternative pathways that make sustained life possible.

Other examples of new biotechnological developments include the recent description of a ganglioside associated with neuroblastoma. This ganglioside is a sialic acid containing glycosphingolipid and is located mainly on cell surfaces. It can be detected in the blood only in the presence of neuroblastoma - and cannot be detected in remission or in normals. It has been identified as being identical to the ganglioside G from Tay-Sachs disease and can be identified by enzymatic degradation with thin layer chromatography. If these observations are confirmed, it appears that this tumour antigen can be used not only for the diagnosis of this particular cancer, but also in the assessment of therapeutic response. The ideal tumour marker has been described as being of tumour cell origin that can be detected in the circulation in proportion to tumour cell mass, and is not found in normals.

Another example of this type of development relates more directly to your neurological research interests. A Saskatchewan boy who did his medicine at the University of Toronto and took a Ph.D. at Harvard, became an expert in nucleic acid and protein changes in brain cells related to memory function. Dr. Sam Bogoch described a new protein entity that seemed related to brain cell memory 'activity'. In studying the presence of this substance, he hypothesized that the most 'active' brain cell should be a tumour cell, and indeed found this protein in large quantities. He identified it as 'malignin' and eventually defined it as an undifferentiated cancer antigen found to be present in the blood of many patients with a variety of tumours. After extensive testing, it appears that this 'tumour antigen' is consistently present in many cancers, and circulates in levels proportional to tumour mass - decreasing following surgery or irradiation and disappearing on cure.

Canada now has a number of small developing 'pharmaceutical' companies. There are at least six or eight small diagnostic companies that have been spun off from academic institutions in the monoclonal antibody field alone. One more company, with which I have been associated for a number of years, is Chembiomed Ltd. in Edmonton. About eight years ago, I was consulted by the University of Alberta about a new development by one of its professors, namely the first chemical synthesis of human blood group antigenic determinants. This was before monoclonal antibody technology was available, and their initial interest was in making a superior line of 'human' blood grouping antibodies in goats, using these synthetic antigenic determinants. After Milstein's description of monoclonal antibodies, Chembiomed converted its process accordingly, and has now come up with a range of blood grouping diagnostic products produced with this technology in mice. Indeed, of the first four monoclonal antibody-based blood grouping products approved by the Bureau of Biologicals in Washington, three of these were Chembiomed products.

This development was made possible because of the first-class academic research of Professor R.U. Lemieux. He has now extended his work and is using sophisticated NMR and monoclonal antibody techniques to describe the role of carbohydrates, as glycolipids and glycoproteins, in intercellular communication. This is very important to our better understanding of such diseases as cancer - especially as it is already known that blood group antigens are involved in a variety of ways in the carcinogenic process.

Another development in the biotechnology field that will be of some interest to those who remember Drs. Hoffer and Osmund, is the recent description of the modification of behavioural response to potent antipsychotic drugs by ascorbic acid. Vitamin C administration results in a significant increase in this vitamin's level in the neostriatum - resulting in an increase in neurostriatal firing rate. High doses block the amphetamine response, suggesting an alteration in dopamine-mediated behaviour. This relates back, of course, to the use of megavitamin therapy in schizophrenia and how it might potentiate or alter the action of antipsychotic drugs.

There is also the recent description of an anti-atherosclerotic plaque monoclonal antibody. This MCA complexes specifically with atherosclerotic foci and if the antibody can be selected appropriately, could be expected to activate complement, mobilize macrophages and accelerate the phagocytosis of accumulating lipids. Indeed, it could result in the development of an anti-atherosclerosis 'vaccine'.

I would not be living up to my reputation if I did not say something more specific about patents and the development of biotechnology. Most people believe that the patenting of novel life forms originated with the Chakrabarty patents, but microorganisms have in fact been patented for over 30 years. At the international level, there are a variety of problems

impeding development, and these include reciprocal recognition, conflicting patent procedures between countries, and the observance of these by the various national patent offices. One of the requirements for patenting a life form or microorganism is its deposition with an official culture collection agency. However, especially the Japanese and some Europeans have raised barriers to the free exchange of reference cultures, and are thereby impeding the free flow of biotechnology essential for development of the field on an international scale. In addition, litigation and perceived indecisiveness of the patent courts have introduced a certain element of unpredictability, making it more difficult to make long-term business decisions.

In Canada, we have a particular problem, namely that of compulsory licensing. This is nothing less than 'idiotic', although a majority of the provinces now see this as such and there is a developing consensus that the law should be changed. However, as late as January 1, 1985, the patent for one of the few remaining major products still available in Canada was licensed out--the federal government not waiting for the final report of the Eastman Commission on the pharmaceutical industry.

In summary, and based on your previous extensive discussion of technological change, I would quote Drucker in that, 'it is no longer a question of whether Canadians want technological change, but how to make do with and accommodate to the inevitable--relating it to a period of economic growth with social justice'. In order to do this, we must somehow expose science as an open social value and better integrate it into our socio-political system. And to do this, we need government support--but this must be measured against well-defined political goals. Otherwise, science will fall victim to the bureaucratic process itself--replete with all the problems of bigness. Government support deforms in various ways and we must be forever vigilant of this.

But in closing, I would like to make a series of recommendations. These are based on Canadian society giving greater priority to the free enterprise system and to the provision of an improved profit-oriented industrial climate. Federal government laboratories should now be given a better defined purpose--especially any new facilities. They must somehow be subjected to more direct peer review, with representative oversight committees. Industrial assistance programs should be simplified and should be made more risk-taking and less risk-averse. In the long run, this government assistance must be increasingly provided as tax incentives rather than administered grants. In this way, technological activities will be more directly related to company profits and losses.

Without question, there must be increased basic research support for universities. This should include greater, if not full coverage of research overhead. There should be a major coordinating role for NSERC and MRC in the funding of university research. In addition, there should be special incentives for universities collaborating with industry. Federal funding mechanisms must be used in some way to encourage the development of specialized academic institutions in Canada.

Canada must take better advantage of strategic government procurement and this will be increasingly based on multi-country joint-ventures. Most important problems of the day are international in scope and these can only be fully satisfied through international production sharing. We must open up our markets to international competition in order to make our industry more competitive.

Compulsory licensing with regard to pharmaceutical and food products should be abolished and the Patent Office moved out of the Department of Consumer and Corporate Affairs. The Exchequer Court of Canada must be reformed, in that their decisions of the last 10 to 15 years relative to industrial property rights have been technologically inappropriate and very detrimental to the development of high-technology industry. Canada must codify plant breeder's rights, if we are to take full advantage of the opportunities of biotechnology in agriculture.

Various methods must be tried to improve the technology transfer process. Specialized institutions such as the National Research Council's National Biotechnology Research Centre, Alberta Research Council laboratories in Edmonton, Ontario Research Foundation and the University of Waterloo Biotechnology Centre should be developed further. Canadians must foster the study of industrial innovation or entrepreneurship in their academic institutions. Perhaps Canada needs a NASA type of specialized effort for biotechnology? An investors' fund could be organized - by asking the top 200 industrial companies in Canada to contribute 1/100 of 1% of their pretax profits, and the life insurance industry, which has not contributed to new health care innovations in this country, might be asked to underwrite the expenses of an inventors' foundation for the first five years.

Canada should consider the institution of cash awards for inventors, and we should consider a museum for 'simultaneous invention'. As has often been pointed out, major developments are almost never alone - the same or related ideas often arising in more than one country at or about the same time. Canada could use the development of a simultaneous invention museum to highlight its own activities and also contribute to world-wide technology development. In addition, there should be a scholarly study of the state's role in industrial innovation, for Canada must learn from its experience; otherwise it is condemned to repeat history.

However, most of all, we must somehow rediscover in Canada the 'free enterprise system', if we are to benefit fully from new industrial developments--the intellectually free as well as economically free enterprise system.

DISCUSSION OF PART II

J.M. CLEGHORN, M.D.

Brain Imaging with PET - A Scientific Bridge

I suggest that chemical and structural localization of function and dysfunction in the brain using positron emission tomography provides a scientific bridge between clinical neurology and psychiatry on one hand and neuropathology and neurobiology on the other. Given the possibility of visualizing and measuring radioactively labelled materials including basic substrates such as oxygen and glucose, it should be possible to construct pictures of metabolic processes in mental disorders. In addition, the labelling of one neurotransmitter, dopamine, and several neuroreceptors, i.e. opiate, benzodiazepine, one neuroleptic and some anti-convulsant drug neuroreceptors means that it will become possible to measure abnormalities in neurotransmitter systems in the brain and to visualize their location.

I have a question for the clinical researchers. Are clinical research methods adequate to meet the challenge provided by remarkable advances in technology? I suggest two answers; yes and no. Yes, if we study well known disease models, and no or questionably, when we have to categorize patients by conventional psychiatric diagnoses. I will try to explain.

Study of the psychiatric aspects of the following several diseases should be profitable.

1. Parkinson's disease is associated with depression of mood and with a memory disturbance that is often ignored. These can be measured. Fluorodopa concentration can be measured also and related to these dysfunctions. Chemical concentration and anatomical localization should be specifiable.

2. In Alzheimer's the reductions in glucose uptake have already been described to some extent. This will provide an opportunity for testing the efficacy of treatments. The acetylcholine receptors have yet to be labelled and measured by positron emission tomography, but when possible, a rare opportunity to assess the possible efficacy of treatments will become available.

3. In Huntington's disease, a psychosis may be the first manifestation and this may occur before structural damage is visible on the CT scan. Rather the PET scan shows reduced uptake of glucose in the basal ganglia. This is a possible model for a type of psychosis and the role of the basal ganglia can be assessed.

4. Multiple Sclerosis, in its fairly advanced stages, often involves what looks like a frontal lobe syndrome, even though there are no frontal lobe lesions. A state of indifference, jocularity, poor judgement and impaired social relations are characteristic. This

appears to be due to disconnection of frontal-limbic tracts due to MS lesions. Their location on PET scan should be specifiable and this important component of a number of mental illnesses studied according to this model.

5. Temporal lobe epilepsy simulates a wide variety of psychiatric disorders -- hallucinations, delusions, and numerous personality changes. These can be examined by studying the alterations in glucose uptake and the relation to the clinical picture, correlating sites in the temporal lobes with clinical manifestations.

6. It will be profitable to compare Tourette's disease and schizophrenia. Both of these conditions apparently involve dopamine supersensitivity and both of them respond positively to neuroleptics and yet they bear no clinical similarity to one another. Localization of fluorodopa concentration or dopamine receptor concentration in these two conditions ought to help localize the sites of disorder in both of these conditions.

When it comes to interpreting the findings in major psychiatric disorder, our reliance on diagnostic categories does not serve us well. We may have to look at discrete phenomena within diagnostic categories. For example, it has already been shown that chronic schizophrenia and affective disorder are associated with a reduced anterior posterior gradient in glucose uptake. This is apparently due to increased glucose uptake in the posterior part of the brain.

If we examine acute schizophrenic patients, this finding obtains in low level slices but not in mid-frontal lobe slices where, if anything, increased frontal lobe glucose uptake is observed. Furthermore, initial findings of lateralization of one frontal lobe over the other in glucose uptake disappears when the patient is examined on the second day. There appears to be an order effect. Thus, the interpretation of findings in major mental illness requires a different approach, one that must be based on a knowledge of functional neuroanatomy of the brain; the psychological function of different structures and the localization of different biochemicals. The new functional neuroanatomy is being developed with remarkable speed at the Philadelphia Center and at UCLA. In time it will be necessary to localize more sophisticated functions than have traditionally been assessed by neuropsychological testing, including for example, the planning of tasks and discourse of speech and the recognition of facial expressions. Such features are relevant to deficits in schizophrenia and the locations of these functions and abnormalities need to be identified anatomically and validated anatomically by the study of rare cases in which lesions occur in specific areas and result in deficits that can be demonstrated on testing. This kind of validation of psychological tests is heretofore virtually unknown in psychiatry, except in well established neuropsychological functions. Given a functional neuroanatomy and chemical anatomy of the brain which can be visualized on PET scan and quantified, our knowledge will be enriched by strategies of study which permit us to give

fairly specific test challenges of a psychological nature. In addition, pharmacological challenges can also be employed. Drugs that are agonist and antagonist to known neurotransmitters and possibly those influencing peptides (when it is possible to label and visualize them) will open a huge possibility for study. The specificity of both drugs and psychological tests though, is admittedly less than perfect and will impose limitations on the studies that can be done. However, that does not mean that a great deal cannot be learned with the techniques that are presently available and have not yet been applied to images of the brain in severe psychopathology. At the same time, correlations between neuropsychological and pathological assessment on the one hand, and structural and chemical disorder on the other may give sufficient information to permit the development of animal models that have been so elusive in the study of psychopathology.

Finally, to answer the question, are clinical methods adequate to meet the challenge provided by high technology, I think it is clear that a great deal can be done but it may be that the real limiting step in all of this is our continued difficulty with the phenomenology of psychopathological states and the need for a more sophisticated neuropsychology of psychiatric disorders.

PART III. SOCIO-ECONOMIC IMPLICATIONS AND HUMAN ASPECTS

ECONOMIC EVALUATIONS OF THE TREATMENT OF MENTAL ILLNESS

PRAN MANGA, Ph.D.
University of Ottawa

Benefit-cost analysis, cost-effectiveness analysis and cost-utility analysis are alternative forms of economic evaluations that have been increasingly utilized to make a variety of decisions in the health care system. One or the other techniques of evaluation have been used to address three different types of questions. Perhaps the most common is the application of cost-benefit analysis to disease-specific programs of intervention such as heart disease, syphilis, polio and kidney disease. A common focus of such analysis is on the economic impact of premature death and disability, the assumption being that the opportunity cost of illness is measurable by the age and sex-specific average earnings of the patients being treated. Benefit cost analysis and cost-effectiveness analysis have often been used to assess which of the alternative treatment procedures or interventions is preferred to attain a given health outcome or status. A noted application of such analyses is the attempts to determine the optimal mix of interventions and regulations to reduce motor vehicle deaths and injuries. Finally, benefit-cost analyses have often been employed to calculate the return on public investment in a variety of health related activities such as manpower training and biomedical research.

These techniques have, of course, been used in evaluating treatment alternatives for mental illness, but "the scarcity of literature in this area is astonishing" (Greene, 1983, 207). The situation is however, likely to change for a number of reasons. There is a marked increase in the incidence of mental illness and its diagnosis and treatment. Governments are the principal source of financial support and they are increasingly obsessed with fiscal restraint and anxious to improve the efficiency of their health care delivery systems. The effectiveness of the treatment of mental illness has been and remains very difficult to assess. A substantial increase in the numbers and variety of mental illnesses makes professionals raise questions about the potential for manpower substitution and alternative modes of treatment. The proliferation of allied mental health professionals has led to a blurring of the proper domain and functions of the various professions, especially in public inpatient and outpatient facilities. The growth was primarily a consequence of a rapid increase in mental illness and the relative unwillingness or inability of psychiatrists to treat certain diagnostic types (for example, alcoholics), patients from the lower socio-economic classes and the elderly. (Redlich and Kellert, 1978). Innovations in drug therapy and the deinstitutionalization movement has provided further impetus to experiment with and assess alternative delivery systems. For all these reasons and no doubt others as well, it is likely that these evaluation techniques will be relied upon more frequently to "guide" public policy decision-making (Saxe, 1980). The word "guide" is used advisedly. The burden of much of this presentation is to argue that it would be both wrong and presumptuous to pretend that any of the economic evaluation techniques can usually provide public policy-makers or their

surrogates with sufficient justification to make a particular decision. While such analysis can be helpful "it seems well to recognize that such analysis cannot hope to eliminate the need for judgement. It can structure the debate, point up the choices and their implications, and various disagreements. But in the end, benefit-cost analysis is an aid to policy formulation - not a substitute for it" (Weisbrod, 1983, 170).

In what follows, the important and common difficulties that one encounters in such analyses are presented. The discussion will highlight those analytical deficiencies that are of critical importance and relevance to decision-making regarding the treatment of mental illness. It will be suggested that the area of mental illness presents particularly complex and difficult issues that are bound to bedevil the analyst irrespective of the evaluation technique that is selected to assess the treatment or policy options. It will be argued that such analyses are invariably incomplete and partial, and cannot avoid or get around important aspects of benefit and/or costs that are practically immeasurable, at least in the terms that need to be employed in such analysis. All this is not to suggest that such analyses offer no guidance to decision-making, but only to point out the limitations and the partial utility of such evaluation techniques.

Almost all of the public policy recommendations derivative of these economic evaluations have been essentially of two types. The first deals with whether or not to start a new health program or treatment to address a particular health problem. The second addresses the question of expanding existing programs in order to achieve the benefits with a particular treatment or diagnostic strategy or, conversely, to contract one that is "proven" to be less effective or efficient than previously thought. These evaluation techniques have not been employed to address the qualitative rather than purely quantitative aspect of public policy as it relates to health. The deficiency is particularly noteworthy when making public policies about mental illness.

It is increasingly appreciated that the economic evaluation techniques are not value-neutral and that, in fact, they either explicitly or implicitly embody some ethical principles or positions. An attempt will be made to highlight the ethical issues that arise in decision-making regarding treatment alternatives in mental illness. It is, of course, not the intent of the paper to address all or most of the important ethical issues that arise in the context of public policy decision-making regarding mental illness. Rather the objective is much more limited. The ethical issues that will be discussed are those that are particularly germane to the economic evaluations mentioned above or those that need to be addressed if and when decisions regarding mental illness appear to be based on such analysis. It is often said, and quite correctly, that cost-benefit analysis and cost-effectiveness analysis are typically concerned only with questions of economic efficiency to the neglect of other determinants of public policy. Such a perspective will inevitably be considered too narrow and insufficient, especially with respect to mental illness. It is hoped that the discussion on the ethical issues presented in this paper will broaden the scope of factors that must be taken into account in making decisions in this area of health care.

Problems encountered in the economic evaluations of the treatment of mental illness

In this section the difficulties and shortcomings of using cost-benefit, cost-effectiveness or cost-utility analysis in evaluating mental illness treatment alternatives are presented. The difficulties are related to a) the theoretical and methodological bases of these analytical techniques, b) the choice of the objective function implied in the analysis c) the lack of analysis of alternative treatments to provide for an appropriate counterfactual, d) issues related to the measurement and outcome conceptualization specific to the treatment of mental illness, and e) the uncertainties and special estimation difficulties encountered because of the varying time paths and duration of the cost and/or benefits of treatment.

The techniques are what economists in their more modest days referred to as "third-best" approaches to economic evaluations. What was emphasized is the partial framework and the fundamental micro-economic orientation of these analytical techniques. They were valid and appropriate to use as long as the effects of the "project" (in our case "treatment options") had only marginal impacts on the economic, social and demographic structures of a society. One of the underlying assumptions of these techniques is that the prices, or quantities of the supply of human and capital resources, or, in the pattern of demand for a given set of health services, will not be significantly altered as a consequence of a particular treatment approach. Fortunately, these basic microstructural assumptions are not violated when one is considering a great number of alternative treatments of mental illness (e.g. alternative drug therapies and psychotherapy). However, some treatment alternatives do have a potential for a multiplicity of interactions and extra-marginal effects on the economic, social and demographic structures of society. A benefit-cost or cost-effectiveness framework may not be appropriate if one was to analyse a major breakthrough in the organization and delivery of mental health services. For example, whether the deinstitutionalization movement can be construed to represent treatment options with significant extra-marginal effects is an important question. An affirmative answer is equivalent to a questioning if not the rejection of the simple application of cost-benefit or cost-effectiveness analysis.

There is an important corollary of this point of special relevance to the evaluation of mental illness treatment. Given the highly social nature of mental health theorizing, assumptions and public policies, analysis that focus only on the patients are an incomplete and inadequate guide to the determination of public policies. The need to abandon the individual patient models to a more comprehensive systemic approach to evaluation has, of course, a long history in the mental health literature (Mechanic, 1969; Wing and Brown, 1970; Wing and Hailey, 1972). In the jargon of economists, the mental health field displays extensive and substantial "externalities". The dispersion of the economic and psychic costs and benefits

beyond the patients themselves is thus likely to be substantial in the mental health area. Any economic evaluation of a project or program of treatment should therefore consider all the consequences extending beyond the target patient group.

In economic evaluations of alternative treatments of mental illness one is likely to encounter a maddening range of methodological and conceptual difficulties. For example, not all of the costs of treating mental illness fall within the health care delivery system. It is common for analysts to ignore costs that fall on patients themselves, and sometimes more importantly, on their families and other public sector agencies. Also, not all costs involve expenditure; for example, the efforts of family members, friends and volunteers. Conversely, not all of the incurred expenditures involve real resource cost to society, that is, some expenditures are merely transfer payments and do not involve the use of real human and capital resources.

There are, of course, further difficulties. Conceptually, it is the marginal benefits and costs that are relevant in cost-benefit or cost-effectiveness analysis. Yet the most readily available and used data will almost certainly be the averages. Furthermore, the averages tend to be derived from accounting records which inevitably generate biased estimates of the true social averages. For example, the opportunity cost of land and depreciation of public buildings is often ignored or artificially low in the readily available information on costs.

It is often, and sometimes correctly, thought that the costs of treatment are generally much easier to define and measure than the benefit side of cost-benefit or cost-effectiveness analysis. This is especially true in the case of mental illness. In fact, the very definition of mental illness is controversial and the measurement of the "benefits" of treatment is correspondingly vexing. Indeed, the conceptual foundation of measuring improvements in mental health may be such as to make a pecuniary (dollar) measure impossible. Monetary values based on employment and productivity gains are just one dimension of the change in mental health status of patients and are obviously incomplete as a measure of the benefits of treatment. Also, willingness to pay (for treatment) by patients, their families and the government is clearly an inadequate measure because of patients' and their families' inability to make informed choices and because of the externalities referred to earlier. In response to such difficulties, analyses are likely to incorporate several different definitions and measures of benefits. For example, in the much-referred to benefit-cost analysis of a controlled experiment in the treatment of the mentally ill (the Mendota experiment) benefits were measured in terms of a) labour market earnings, employment stability and evidence of financial planning b) clinical symptomatology measures such as motor agitation, depressed mood, paranoid behaviour and c) changes in mental health status as expressed by patients themselves (Weisbrod, 1983).

Analysts often resort to cost-effectiveness analysis whenever they cannot surmount the difficulties in measuring or estimating the benefits of a health care intervention or program. It too is a micro-oriented

technique of analysis. It offers greater flexibility than benefit-cost analysis in that it is not restricted to reducing the effects of medical treatment or value.

Clinicians, of course, regard the change in the patient's mental health as the most important measure of the effectiveness and outcome of the treatment of their patients. The patient's ability to return to work is, of course, one measure of a change in patient welfare but it need not be the primary objective of treatment. In the case of many patients such as the elderly, employment or employability is literally of no consequence with respect to the choice of treatment. As well, the treatment of the mentally ill may confer benefits to the family and colleagues of the patients which may not be reflected fully or sometimes not at all in the patients' productivity and earnings. Indeed, in the treatment of the elderly with Alzheimer's disease, senile dementia and depression, it is the value placed by others on the benefit of treatment that is the dominant component of determining what type of treatment is given and how much care is offered.

In cost-effectiveness analysis it is thus possible to estimate the cost per "unit of effect" and also to disaggregate the analysis on an effect-specific basis. Thus cost-effectiveness analysis could permit both qualitative and quantitative indicators of the effect of treating mental illness. The technique is especially useful where the analysis is to determine the preferred treatment option for obtaining a prior set of agreed upon effects. Indeed, the analysis in such a situation is basically a search for the cost minimizing strategy.

While cost-effectiveness may be thought to be more applicable to evaluation of mental health programs, it suffers from two rather serious analytical disadvantages. First, it is difficult to apply discounting techniques to a flow of effects of the alternative treatment of mental illness over time. The difficulty can be highlighted by way of a question. How should we choose between two treatment alternatives, one of which yielded a relatively small benefit in the reduction of mental illness within the first year, and the other, a relatively larger one but which becomes manifest only five to ten years from the time treatment begins (assuming, for simplification, that the present value of the social costs of treatment were the same)? The matter becomes downright intractable if one assumes the more likely event that the social costs of the two alternatives are also very different, though the treatment with the larger but distant positive effect is also somewhat costlier than the first. It might be noted that such time-paths of costs and benefits are rather likely in the treatment of mental illness. Without employing discounting techniques and common measures of effects it is difficult to make meaningful comparisons between treatment alternatives. "Such difficulties would be heightened where the set of effects of each alternative being analyzed include qualitative indicators as well as quantitative ones. For example, changes in the affective rather than a cognitive learning with respect to substance abuse (drugs, alcohol, etc.) which have health status impact, may be the objective effects of a particular preventive mental health program. At the present time, societies throughout the world have difficulty discussing

such issues rationally, let alone measuring, with any degree of accuracy, effects of such programs, taking into consideration the time-path of those effects" (Dunlop, 1980, 356).

The second difficulty with cost-effectiveness analysis arises when the alternative treatments of mental illness have differing sets of possible effects. Cost-effectiveness is appropriately employed in those situations where the effects of alternative treatments are similar. The problem is, of course, that in the mental health area, different treatments of mental illness often yield a different set of effects. Thus while cost effectiveness analysis avoids the need to establish a common monetary measure of benefits it often presents the problem of having to develop a "numeraire" from dispersed or differentiated measures of benefits.

A brief comment on cost-utility analysis is appropriate at this point. In all three forms of economic evaluation the costs are measured in much the same manner. It is the measurement of benefits or outcomes that differentiates them. Whereas the benefits under cost-benefit analysis are measured in monetary terms, they are measured in terms of clinical, psychological or general mental health status terms under cost-effectiveness analysis. Under cost-utility analysis, these health status benefits are transformed into a utility measure, for instance, the improvement in the quality of life resulting from the effects of the treatment of mental illness. From the foregoing discussion, it should be apparent that, in essence, cost-utility analysis attempts to overcome the second of the two difficulties facing cost-effectiveness analysis. It is equally apparent that the application of such analysis in assessing alternative treatment of mental illness would be very difficult and controversial.

The problems generated by the different time-paths of the benefits (however measured) and costs of treating mental illness are not to be minimized. They have to be taken into account if any economic appraisal is to be useful for purposes of public policy. The technical problems are well-known and will therefore only be mentioned here.

In economic evaluations the costs and benefits that occur in the future must be regarded as expected and by no means certain (Saxe, 1980). It is obvious that there are risks or uncertainty associated with the future effects of treating mental illness. A good illustration of the importance of taking a longer term view is offered by a study by Pasamanick et al (1967). They compared the effects of hospitalization and home treatment on several groups of schizophrenics observed for 6 to 30 months. The authors concluded that home care was feasible and effective in preventing hospitalization. While the study can be severely faulted for ignoring a number of social costs what is interesting is that most of the improvement in the patient's mental health and social functioning occurred during the first six months. There was very little, if any, improvement thereafter. Thus one locus of treatment appeared to have no distinct advantage over the other, at least in this patient group at that time. A five year follow-up of these patients revealed no difference in social and mental functioning between the hospitalized and the community based patients.

Another difficulty is the issue of the appropriate discount rate to use in comparing the costs and benefits over time. Both discounting and the choice of the discount rate can make a significant difference to the final evaluation (Drummond, 1980). As a practical matter, several alternative discount rates can be employed to undertake a "sensitivity analysis" of the results to determine the extent to which the ranking of the alternative treatments is altered because of the variations in the discount rate.

The last point brings up yet another problem encountered in these analyses. It is the problem of determining the appropriate measure to employ in ranking alternatives. Should one calculate and rank the alternative on the basis of net present value, benefit-cost ratio, cost-effectiveness ratio, internal rate of return, or whatever other criterion one may devise? The extensive literature on these evaluation techniques suggests that the net present value is the most appropriate ranking criterion (Drummond, 1981). It is the most problem-free decision index, even though it is often not easy for public policy-makers to interpret its significance. In the area of mental illness, however, the following general caveats are of considerable importance. The criteria used in economic appraisals should not be viewed as hard and fast rules of decision-making in ranking treatment options. Such criteria are merely summary indicators of the results of economic evaluations, results which in the nature of these analyses in the mental health area are inescapably incomplete and subject to considerable margins of error in the estimates of benefits and costs if and when such estimates are feasible.

The notion of ranking alternatives obviously presumes that for economic evaluations there must be at least two alternatives to appraise. Indeed, even when just one specific treatment is being assessed, the alternative of "doing nothing" requires explicit analysis. That is, the treatment program is evaluated against the conterfactual of no program. An important corollary of this requirement is that the relevant comparison in evaluation is "with" versus "without" a particular treatment, not "before" versus "after" the treatment. Alternatively, the proper comparison is between one program versus another or others. It is the existence of alternatives that gives the scope and legitimacy for economic evaluations.

The conceptual importance of the notion of a counterfactual should not be underestimated (McGuire and Weisbrod, 1981). In fact, it is quite central to the determination of just what are the relevant costs and benefits that should be incorporated in an economic appraisal. For example, treating mentally ill patients in hospitals involves "costs" of providing food. Were the patients to be "deinstitutionalized" (that is, the counterfactual of the mental hospital) they would still have to eat. If both treatment modes involve the same food cost, there is no effect on the experimental program deriving from that particular cost variable. "This principle of including in costs or benefits only those consequences that would not occur in the absence of the program is of fundamental importance in program evaluation. It underscores the need for an explicit statement of the alternative with which the program being evaluated is being compared". (Weisbrod, 1983, 44).

Perhaps the most valuable research approach that provides for an acceptable specification of a counterfactual is the randomized control experimental design. The ethical and political problems of withholding "effective" - usually assumed rather than proven - therapy often vitiates such experiments (Kamper-Jorgensen, 1977; Zusman and Bisonette, 1973). Alternatively, "natural experiments" in which "similar" population groups were subjected to different treatments could provide the desired counterfactual. One difficulty with such analyses arises from "selection bias", that is, patients often select themselves to receive a particular treatment or have been selected to receive it by their physicians or family.

Another and perhaps more complex problem with natural experiments is that usually more variables differentiate the two or more alternative treatments than the analyst can deal with systematically. For example, in a number of analyses comparing the effectiveness of a treatment in an earlier period to another treatment prevailing at a later period, the analysis ignored changes in one or another aspect of care that could materially affect the results, for example, the increased availability and use of antipsychotic drugs over time (May, 1970). It is easy to mistake the effects of a new therapy for effects of other variables that also influence mental health.

In the field of mental illness, it is not uncommon for there to be no counterfactual, and if there is, to have just one. Also, the treatment alternatives being evaluated are "composites", that is, many variables are changed simultaneously (for example, hospitalization versus community based mental health programs).

This is not to suggest that the existing cost-effectiveness analyses of the treatment of mental illness do not involve more than one alternative, but only to point out that there are very few such studies. One example of such a study is by May (1971) in which several treatment options for first-admission schizophrenic patients were appraised. He found that drug therapy and electro-shock therapy were the least costly treatment options and that psychotherapy was the most costly form of treatment per patient released. However, the study made no attempt to assess the long term effects of the treatment options.

While the foregoing discussion only highlights rather than provide an exhaustive listing of the many problems encountered in undertaking economic evaluation of the treatment of mental illness, it by no means makes the case that no such analysis should be attempted. In light of the many and serious difficulties cited this may sound incongruous. But this is a matter of perspective. Benefit-cost analysis and its related techniques must be seen in a more modest role. The techniques are useful in providing a conceptual framework for at least enumerating the possible benefits and cost in the form of a balance sheet, which can then be considered in both quantitative and qualitative ways in making decisions. When thought of as a framework, they are useful in systematizing the decision-making process and form a basis for determining the need for acquiring additional information in order to more fully appreciate the relative merits and demerits of treatment options.

An area that remains highly controversial and where better evaluation would certainly be of great value is the issue of treating the mentally ill at home and in the community instead of the traditional mental hospital. It is the official policy of many governments and has the enthusiastic sanction from professional and consumer organizations. There are claims and counter-claims as to the wisdom of this policy. From the many earlier examinations critical of deinstitutionalization (Brown et al, 1968; Rutter, 1966; Freeman and Simmons, 1963) much indeed was learned so that improvements to the programs could be implemented. The introduction of pharmacologic agents has improved the effectiveness of community care. While community based care can be modified and improved with successive and careful evaluation the history of deinstitutionalization suggests that the policy was implemented without much planning and evaluation. A major shortcoming was the improper and indeed a lack of consideration of alternative modes of delivery and treatment. According to one analyst "the range and sequence of treatment modalities initially seen as offering great hope and promise - smaller, better-staffed hospitals, halfway houses, sheltered workshops, emergency protective resources, and community treatment centers - never were implemented, as political enthusiasm, fed by inflated rhetoric, moved to community treatment, evaluation of social ills, and the elimination of publicly supported mental institutions. It is highly conjectural that the severely mentally ill have had their lot that much improved in the process" (Arnhoff, 1975, 1281). The excesses of deinstitutionalization were of course corrected in many instances though problems remain. The history of deinstitutionalization, certainly in its earlier period if not at present, raises questions about the bases for major policy changes. One question that needs to be carefully assessed is whether the economic and fiscal needs of the state are regulating and redirecting the medical determination of the best approach to care and treatment (Rose, 1979). In a recent review of mental health policy in the United States, Greene (1983) concluded that "probably more than any other array of health services, the structure of the mental health system has been defined by economic and financing issues. Understanding it, to a very large degree, rests with the analysis of its administrative history and an explanation of its administrative requirements" (p. 212).

Ethical Aspects of Evaluating the Treatment of Mental Illness

It is commonplace to acknowledge that economic evaluations rarely take into account ethical aspects of the treatment of mental illness. This is understandable since many, though not all of the ethical issues are beyond economic assessment and consideration. Psychopharmacology has had dramatic effects upon a number of mental illnesses (for example, schizophrenia) but one may argue about the ethics of controlling a person's behaviour by potent drugs and the risks of iatrogenic neurological sequellae that could result because of the long-term and sometimes haphazard and massive dosage used to prevent rehospitalization. Tranquilisers are, it is often said, among the most overused, misused and abused drugs in the pharmacopoeia. The subjection of patients to massive doses of phenothiazines in hospitals to subdue violent or aggressive behaviour also involves ethical questions. Similarly for the use of electroshock therapy. These

issues do not, however, involve only clinical matters in the choice of intervention, (that is, how mental illness is treated or how behaviour is changed). They also raise questions about for whose benefit the particular therapy is chosen or behaviour control obtained (London, 1973).

Psychosurgery is a dramatic, potentially dangerous and highly controversial therapy (Fromer, 1981, 240-243). It is considered an experimental procedure, to be performed only when medically indicated and when the patient has given informed consent. Any economic evaluation of this form of therapy will, of course, be highly problematic. The danger and risks of psychosurgery involves the individual patient and his family members but not the larger society. How does one assess a patient's altered identity after surgery? Does the larger society's valuation of the patient's altered state matter in the analysis? These are obviously difficult and provocative questions.

The underlying issue that has to be resolved in addressing these questions is whose goals the therapy is directed toward achieving - the patient's, his or her family, the therapist's, or society's. Different objectives will yield varying assessment of the merit of such dramatic treatments.

It is sometimes the case that a therapy may be both efficient, in the economists' sense of the word, and efficacious, in the sense that the therapy produces the desired clinical or behavioural result. Yet the ethics of the therapy would be highly questionable. It is at least arguable that it may be difficult to justify aversive therapy even in the case of a "child molester" when painful and dangerous stimuli are used to achieve a therapeutic goal. Such cases should remind us that even if the economic evaluations were thorough, complete and unambiguous, public policy makers must examine other aspects of treatment in assessing the advisability of adopting a particular treatment option.

The evaluative technique typically used in assessing alternative diagnostic or therapeutic intervention in illnesses concentrate on the ends of the intervention, that is, on questions of efficacy, outcomes, efficiency and effectiveness. However, over the past two decades much concern has been focussed on the means of treatment. It is not enough that therapy brings about the results the professions or society may desire. Indeed, many of the patient rights, such as the right to refuse treatment, the right to be informed about the benefits and risks of treatment, the right to receive care in the least restrictive alternative treatment setting, can be construed as protecting the patient, not only from professions but from society at large and its government and agencies.

While these are serious questions virtually ignored in most evaluative research there is one ethical issue that is often considered or at least acknowledged to be relevant in cost-benefit or cost-effectiveness analysis. Public policy is not always or, perhaps more accurately, not only interested in questions of efficiency or efficacy but also in issues

of equity and distributive justice. To address such concerns methodological improvements in the framework of benefit-cost analysis and cost-effectiveness analysis are beginning to be made, though more often in theory than in practice. The general approach is to disaggregate the analysis of the alternative treatment and to conduct it on a group-specific basis including age, sex, and income or socio-economic status. The benefits and costs of many treatment options are differentially borne by patients and their families. Community based mental health services and the deinstitutionalization policy generally have rarely been assessed for their distributional impacts. Assertions that the poor do not receive sufficient care or are less likely to be treated or receive poorer quality of care are some examples of the distributional issues raised in the literature. Similarly the general prejudice and misconception that the elderly are beyond psychiatric treatment or care has often meant woeful neglect of the mental health problems of the elderly (Hall, 1982) while it may be more "efficient" to devote more resources to the younger generations, consideration of social justice may suggest that we allocate increasing amounts of professional and technical resources to the elderly.

To formally assess such equity issues, it is imperative that more comprehensive and detailed economic evaluations be attempted.

REFERENCES

Arnhoff, F.N., "Social Consequences of Policy Toward Mental Illness," Science 1975, vol. 188, pp. 1277-82.

Brown, G.W., M. Bone, B. Dalison, and J.K. Wing Schizophrenia and Social Care: A comparative follow-up of 339 schizophrenic patients, Oxford University Press, London, 1968.

Brown, M.F. Studies in economic appraisal in health care, Oxford University Press, London, 1968.

Drummond, M.F. Principles of Economic Appraisal in Health Care, Oxford University Press, Oxford, 1981.

Dunlop, D.W. "Benefit-Cost Analysis: A Review of its Applicability in Policy Analysis for Delivering Health Services" in S.J. Williams (editor) Issues in Health Services, John Wiley and Sons, New York, 1980, pp. 351-362.

Freeman, H. and O. Simmons, The Mental Patient Comes Home, Wiley, New York, 1963.

Fromer, M.J. Ethical Issues in Health Care, C.V. Mosby Company, St. Louis 1981.

Greene, B.R. "Evolving Mental Health Policy: Implications for Management Development in Health Administration", The Journal of Health Administration Education 1983, Vol. 2, No. 2, pp. 193-220.

Hall, M.J. "Mental Illness and the Elderly" in R.J. Vogel and H. Palmer (eds.) Long Term Care, U.S. Department of Health and Human Services, Washington, D.C. 1982, pp. 483-506.

Kamper-Jorgensen "Scientific Methods in Long-Term Outcome Evaluation of Preventive Health Programs: A Critique of the Randomized Controlled Trial" Scandinavian Journal of Social Medicine, Supplement 13, 1977, pp. 81-91.

London, P. "The Ethics of Behaviour Control" in R.W. Wertz (editor) Readings on Ethical and Social Issues in Biomedicine, Prentice Hall, Inc., New Jersey, 1973, pp. 162-178.

May, P. "Cost-efficiency of Mental Health Delivery Systems" American Journal of Public Health, vol. 60, 1970, pp. 2060-67.

May, P. "Cost-efficiency of treatments for the schizophrenic patient" American Journal of Psychiatry, vol. 127, 1971, pp. 1382-5.

McGuire, I.G. and Weisbrod, B.A. Economics and Mental Health U.S. Government Printing Office, Washington, D.C. 1981.

Mechanic, D. Mental Health and Social Policy, Prentice-Hall, Englewood Cliff, N.J., 1969.

Pasamanick, G., Scarpitti, F.R. and Dinitz, S. Schizophrenics in the Community, Appleton-Cenetury-Croft, New York, 1967.

Redlich, F. and S.R. Kellert, "Trends in American Mental Health" The American Journal of Psychiatry, 1978, vol. 135, pp 22-28.

Rose, S.M. "Deciphering Deinstitutionalization: Complexities in Policy and Program Analysis" Milbank Memorial Fund Quarterly, vol. 57, 1979.

Rutter, M.K. Children of Sick Parents: An Environmental and Psychiatric Study, Oxford University Press, New York, 1966.

Saxe, L. The Implications of Cost-Effectiveness Analysis of Medical Technology, Background Paper no. 3: The Efficacy of Cost-Effectiveness of Psychotherapy" Washington, D.C. Government Printing Office, 1980.

Wing, J.K. and Brown, G.W. Institutionalism and Schizophrenia, Cambridge University press, Cambridge, 1970.

Wing, J.K. and Hailey, A.M. (eds.) Evaluating a Community Psychiatric Service: The Camberwell Register 1964-7, Oxford University Press, London, 1972.

BIOTECHNOLOGICAL ADVANCES AND THE DESIGN OF MENTAL HEALTH SYSTEMS

MICHAEL DEAR, Ph.D.
McMaster University, Hamilton

ABSTRACT

This paper is a preliminary assessment of recent biotechnological advances in psychiatry as they relate to the design of mental health care systems. A review of a family of planning models suggests four evaluative criteria: availability, utilization, assignment, and distribution. It is suggested that the search for a relevant and appropriate biotechnology is hampered by the current "biotechnological context". These cautionary remarks are not intended to be anti-scientific or anti-progress; but they are decidedly pro-care in their emphases.

My assessment of the human implications of biotechnological advance in psychiatry is inevitably influenced by my recent research in California.[1] This state is only now beginning to experience the full impact of deinstitutionalization. In the early part of 1985, newspaper reports have highlighted the following problems:

- There are at least 34,000 homeless people in Los Angeles County, 40% of whom are ex-psychiatric patients, and 40% of whom have drug or alcohol problems.

- The Los Angeles Police Department intends to put "mental health experts" into police stations in an effort to help the mentally ill "before they commit violent crimes."

- A new scandal has broken out over the quality of care offered in the State of California's 22,000 board and care facilities, which house over 150,000 people.

Such tangible indicators of a system under stress is offset by other news that a private psychiatrist in Beverly Hills is establishing a CATSCAN facility for routine diagnostic practice.

In the context of such a polarization of events, it seems unreasonable and unwise to divorce issues of biotechnology from other aspects of the mental health system or from the wider problems of resource allocation. The mental health care system--in both its research and its service sectors--is properly regarded as a single integrated system. Resource allocation in one of the system's sectors almost inevitably takes away resources from other sectors.

The problem which I wish to address is how to design a mental health care system which includes a biotechnological component. At the core of this issue is the search for what Dr. Boulton (and others) call an "appropriate and relevant" technology. I recognize that some may consider this to be a service delivery issue, which should properly be regarded separately from biotechnology as a research tool. However, I find this distinction difficult to sustain on ethical, theoretical, or clinical grounds. In assessing the impact of biotechnology, one is inevitably led to consider the practical utility of such a large commitment of human, technical and financial resources.

Hence, in what follows, I attempt a preliminary assessment of recent biotechnological advances in psychiatry as they relate to the design of mental health care systems. First, a brief review of a family of planning models is provided. These models suggest four criteria for a subsequent assessment of the new biotechnology. Finally, an alternative model is developed in order to underscore the need for a systemic evaluation of biotechnological advance.

One caveat is necessary. In the space of one short paper, I cannot hope to present a critique which would encompass the wide realm of biotechnological advances summarized in this conference. We have dealt with an incredible variety of diagnostic and therapeutic techniques, new machines, new drugs, and diverse proposals for a research agenda. I am obliged in this discussion to use the term "biotechnology" to encompass all these aspects; inevitably, not all my remarks will apply with equal force to every dimension of biotechnological advance. I must therefore rely on the reader to exercise personal judgement regarding the applicability of my argument to his or her special interests.

1. Models of Health Systems Planning

Most geographical models of health systems planning are based on two simple assumptions: first, that there exists a given configuration of demand for care; and second, that there are available a number of potential sites from which care can be delivered. The planning task may consequently be expressed as an optimization problem. One of the more common mathematical programming models attempts to minimize the sum of fixed costs (i.e., facility capital and operating costs) plus consumer travel costs to obtain care. This so-called "location-allocation" model therefore portrays health planning as a fundamental trade-off between facility costs and consumers' accessibility. A balance is sought between a highly centralized system with low fixed costs but high aggregate travel costs, and a dispersed system with high facility costs but low overall travel costs to the consumer. The model is typically bound by a series of constraints which (for example) provide upper limits on facility and consumer budgets, on the number and capacity of facilities, as well as ensuring that all demand is met (Calvo and Marks, 1973).

A further partitioning of the location - allocation problem reveals a family of related models. These may be termed diffusion, access, assignment, and central place models. I wish to review briefly each of these models, since they provide useful operational criteria by which we may judge the utility of biotechnological advance.

Diffusion models emphasize the rate at which a particular innovation becomes distributed throughout a given region or system. Such models are popular in, for example, the study of the diffusion of epidemics (cf. Cliff and Haggett, 1984). In our case, interest lies in the rate at which a particular resource innovation becomes available in a health care system. We may speculate that the rate of diffusion will vary with many factors, including: the cost of the innovation; consumer demand; profitability of the innovation; amount of advertising; and proximity to the center of innovation.

Models of accessibility are concerned with the utilization of a given resource. Note that mere availability of a resource will not guarantee that it will be accessible or utilized. Utilization in a geographically-dispersed health system has been observed to be a function of three sets of variables (Dear, 1977): service characteristics (including opening times, quality of staff, etc.); client characteristics (including socioeconomic status); and location (including physical proximity, catchment rules, and "social distance").

Assignment models tend to assume the existence of a range of client needs and a set of treatment alternatives. The task then becomes one of matching client needs with appropriate treatment settings (Dear, 1981). The model assumes, usually overoptimistically, that a full range of services exists as well as a set of competent diagnostic routines, and an effective referral machinery.

Finally, central place models recognize the hierarchical distribution of most health resources. The hierarchical principle is a concession to the fact that a full range of services simply cannot be made available at every potential service site. The specialized ("higher-order") services tend to be available in only a very few central locations, such as metropolitan teaching hospitals. The less specialized ("lower-order") services tend to be dispersed throughout the system, as in every neighborhood clinic. The critical requirement in the operation of a hierarchical health system is that referral to higher-order centers has to be easy and efficient (Shannon and Dever, 1974).

This family of related models suggests that we may adopt four criteria in the assessment of biotechnological advance. They are: availability, utilization, assignment, and distribution. These are plainly not the only criteria which could be used. However, they are sufficiently different, in their focus on health care delivery, to provide a useful evaluation matrix for assessing the new biotechnology.

2. Assessing the New Biotechnology

In this section of the paper, I shall employ the concepts of availability, utilization, assignment and distribution to develop a critical assessment of the human dimensions of biotechnological advance in psychiatry. My primary purpose is to generate a range of propositions which may be useful in subsequent analysis.

Availability. Certain innovations are likely to diffuse at a relatively rapid pace throughout a health care system. A new drug, for example, may develop a startling popularity under the combined impact of effective publicity and high potential profitability. Other innovations may be very slow because costs are prohibitive, as in the case of the new imaging techniques (NMR, PET, etc.).

For the moment, the rate of biotechnological innovation diffusion seems to depend upon that most unreliable of scientific criteria--novelty. As one looks "inside the skull" at images of a living brain, it is difficult to avoid a sense of excitement and of imminent discovery. It is also difficult to resist the impulse to leap on the bandwagon of discovery, and to participate in the research program. These desires are especially strong since biotechnology seems to offer a new legitimacy in psychiatry--scientific answers and scientific respectability in a single glamorous technological package.

The appeal of the biotechnological fix remains strong, despite the "unwarranted optimism" which characterized earlier advances (see, for instance, Jacoby, 1981). It is highly unlikely that we shall be able--or should want--to cause a retreat from high technology. But it seems to me that the availability of biotechnology is currently predicated on the two most unstable of theoretical and clinical standards: profitability; and novelty/status.

Utilization. Let us assume that biotechnological advances will continue. It remains likely, however, that (a) their availability will be strictly limited; and (b) resources consumed by biotechnology will not be available for other services (so-called "opportunity costs"). We now face trade-off decisions regarding health care delivery on a scale which has hitherto been unimagined.

First, how do we triage in the biotechnological resources? In the U.S., triage is usually accomplished through selecting only those patients who are able to pay for treatment. In the U.K., triage is by delay--waiting lists regulate the flow of patients, and only the "most needy" are assumed to persevere. Or we may envisage triage by availability, in which case, only residents living near Hamilton or Vancouver would obtain ready access to PET-scan resources. Is it unreasonable to expect triage through a combination of all three methods--ability to pay, delay, and availability? And may we then expect to see lengthening wait lists in biotechnology-rich centers, together with longer length-of-stay in psychiatric units, diminished overall patient loads, etc.?

Secondly, we need to enquire precisely about the opportunity costs of biotechnology. In some ways, we are damned if we do participate in the biotechnological revolution, and damned if we do not. Hence, it has been argued that we must develop an indigenous Canadian biotechnology in order to reap the rewards of the "post industrial society". But even in the U.S. -- the world leader in the development and application of a wide variety of biotechnologies--future profitability and advance is threatened by the relative "lack of funding to develop new engineering technologies related to the production of biotechnology products" (Science, 1984). It is impossible to predict the outcome of the panoply of research and development options in biotechnology. However, we have to be cautious of Faustian bargains which aid a few people at a cost which could bankrupt the entire health care system.

Thirdly, patterns of utilization are, in general, increasingly beset by the trend toward privatization in health and social services (see, inter alia, Metro Toronto Social Planning Council, 1984). More and more services are being "commodified"; that is, being translated from status as a public good (provided by government at public expense) to status as a private good (to be purchased in the marketplace). Recent advances in artificial heart transplants were achieved through a massive allocation of funds and resources to a profit-making medical corporation. Profit-making companies now own or manage over 20% of U.S. hospitals, double the percentage in 1980 (Time, 10 December 1984). At a more mundane, daily level of operation, it may be expected that the profit-making sector is likely to "cream" off only the well-behaved, high-return, curable cases from the patient list. The public sector is then left to deal with the high cost, chronic patient load.

Assignment. If we are concerned to achieve an effective match between client needs and treatment options, then the current performance of many biotechnologies leaves a great deal to be desired. In the first place, many techniques are overprescribed, misused, and of limited diagnostic capacity.[2] In many cases, it appears that expensive procedures are being used to aid a tiny fraction of patient demand. A strong argument could be made in favour of a system which deals with expected needs instead of exceptional needs.

As we refine our tools of measurement, it is as well to recall the limitations of our understanding of the genesis and progress of psychiatric illnesses. Our diagnostic capacities remain limited; thus the probability of affecting an entirely accurate client assignment is small. And even if we can revolutionize our diagnostic techniques, the success of the assignment procedure depends upon the existence of an adequate range of treatment options. Unfortunately, after many years, many regions possess only a rudimentary community-based mental health care system. Moreover, the administrative machinery for client referral is often unwieldy and crude.

Needless to say, an absence of community facilities and poor referral mechanisms can hardly be blamed upon biotechnology. But the facts are: (a) that resources devoted to biotechnological research and clinical practice are resources which are not available to other sectors in the care

system; and (b) that advances in diagnostic capacity are likely to have a minimal impact on client care if the "downstream" components of the delivery system remain underdeveloped.

This having been said, I am happy to concede that only a fool would wish to halt research which shows promise in preventing mental disability (such as the current work on dementias). The prevention of a disease or disability obviously has enormous human and resource implications for the entire health system. The cautionary brake which I am here gently applying to biotechnological optimism is simply intended to raise this issue: if we have limited diagnostic capacity and a largely rudimentary community-based mental health care, what direct benefits can we hope to reap from investment in biotechnology?

Distribution. In most cases, at least for the foreseeable future, biotechnological resources will remain a higher-order psychiatric service. Such resources are expensive, and are playing a significant part in the hyper-inflation of medical costs. The U.S. Office of Technology Assessment recently reported that Medicare costs rose 19% per annum between 1977 and 1982; no less than one-third of this rise was attributed to technology or "intensity" of services (Los Angeles Times, 15 January 1985). Given the relative scarcity of resources, biotechnology will inevitably be rationed through its geographical distribution. At the moment, we seem to be totally without guidelines which would allow us to decide upon the rational distribution of these scarce resources. How many PET-scanning devices do we need? Where should they be located? How do we ensure proper access by patients from remote areas? Who should make the decisions regarding distribution?

3. An Alternative Model: Emphasizing Client Needs

Let us now consider an alternative model for health systems planning. Its implications will allow us to develop a series of tentative propositions which could form the basis for a more formal evaluation of biotechnological advances in psychiatry.

I proceed from the assertion that biotechnological advances should be considered in the context of the total health care delivery system. Mental health care is necessarily provided through an integrated system involving the total range of services -- from psychiatric hospitalization through to home- support for independent living. The task facing us is to ensure that the total variety of client needs is served by an appropriate range of service options. Under these circumstances, the objective function in the optimization routine (see page 4) can be rewritten as follows: to maximize client utilization of the care system, subject to the usual constraints of need, capacity, budget, and so on.

To simplify, this alternative model is designed to maximize the number of clients using the care system. This is not to be regarded as a trivial objective. Its immediate implication (for present purposes, at least) is that technologies and procedures which add only a few clients to the total system utilization are to be resisted. This is especially the

case if the marginal increment in client census is purchased at enormous capital outlay. It is even more to be resisted if long wait-lists for new technologies reduce the capacity of psychiatric units. Overall utilization could even decline under such circumstances. The model suggests an alternative service strategy which would involve the greatest number of clients for the smallest financial investment. (This may be achieved, for example, by giving priority to the development of preventive services).

I shall not pretend that the caricature presented in the last paragraph is a complete and accurate presentation of the problems which beset planning for mental health services. However, I believe that its selective distortion highlights the profound ethical dilemma facing scientists and clinicians alike: What kind of mental health care system do we want? I cannot answer this question here; but I shall provide a summary of the "biotechnological context" in which this question must be considered. I shall present this contextual summary in the form of a series of propositions which can be traced back without difficulty to the preceding arguments in this paper.

1. Biotechnological innovation and availability is promoted by profitability and novelty; it is hindered by high costs.

2. Despite earlier misplaced optimism, it is unlikely that there will be a retreat from high technology; the new legitimacy offered by biotechnology is too powerful to resist.

3. Biotechnological advances will continue to be of limited availability.

4. The opportunity costs of investment in biotechnology are unknown, but are likely to be considerable.

5. Patient triage by delay, ability to pay, and availability are all inherently unfair.

6. New biotechnologies are likely to increase waiting lists and average lengths of hospitalization, and to decrease the patient loads of many psychiatric units.

7. Privatization is a major force in the development of contemporary health and welfare services: it is likely to reinforce the development of biotechnology; at the same time, the potential financial rewards from biotechnology will increase the tendency toward privatization.

8. A significant proportion of biotechnologies are overprescribed and are of limited diagnostic utility.

9. Much expenditure in biotechnology is presently helping only a small fraction of clients in need; the counter-argument that much biotechnology is better regarded as a research tool than a clinical service is fallacious (unless we first answer the question: research for what purpose?).

10. The majority of psychiatric need is not addressed by the new technologies.

11. We currently have no principles which would allow us to decide on the appropriate distribution of biotechnological resources.

12. Integrated mental health systems require a biotechnological component which is relevant and appropriate to the full spectrum of client needs.

13. Who should make decisions on priorities in mental health systems given this highly volatile biotechnological context?

4. Conclusion

This paper has outlined only the preliminary stages in the necessary evaluation of biotechnological advances in the design of mental health systems. The "mirror" which I have held up to examine these advances contains some rather chastening images. I hope they will cause further reflection on the kind of mental health care we want.

It is not too early to begin this assessment. Innovation in psychiatry typically consists of a rapid and widespread adoption of the new idea, followed by a long re-appraisal (which may be literally and metaphorically painful). Only now are we confronting the relative merits of many old technologies, such as ECT.

In advocating this assessment, I realize that I am parting company with at least some of my conference colleagues. For them, the canons of basic science are sufficient motivation to pursue the limits of biotechnology. As this conference has shown, their achievements have been impressive, and their optimism richly rewarded. Hence, I would not wish my cautionary remarks to be regarded as anti-scientific or anti-progress. However, my viewpoint is decidedly pro-care; and I would like the new biotechnology to serve our needs as an appropriate, relevant component of an overall mental health care system.

FOOTNOTES

1 I am grateful for the assistance of Clodagh Stoker-Long in the preparation of this paper, and for the advice offered by McMaster colleagues Dr. B. Humphrey, Dr. G. Stoddart, Dr. D. L. Streiner, and Dr. N. White. None of these individuals is responsible for the opinions expressed in this paper. This paper was written when the author was visiting professor at the University of California in Los Angeles.

2 The literature in this area is very extensive. Some of the most intriguing reassessments have occurred around the use of computerized topography (CT): see, for instance, Jacoby and Levy (1980) and Tsai and Tsuang (1981). D.L. Streiner's essay (in this volume) discusses the limitations of the dexamethazone suppression test (DST).

REFERENCES

Calvo A., Marks D. (1973). Location of Health Care Facilities: an analytic approach. Socio-economic Planning Sciences 7, 407-422.

Cliff A., Haggett P. (1984). Island Epidemics. Scientific American 250 (May), 138-147.

Dear M. (1977). Location factors in the demand for mental health care. Economic Georgraphy 53, 223-240.

Dear M. (1981). Assigning service-dependent elderly to appropriate treatment settings. Social Science and Medicine. 15D, 405-416.

Jacoby R. (1981). Editorial: Dementia, depression and the CT scan. Psychological Medicine 11, 673-676.

Jacoby R., Levy R. (1980) CT scanning and the investigation of dementia: a review. Journal of the Royal Society of Medicine. 73, 366-369.

Los Angeles Times 15 January 1985, part v, p. 6.

Metropolitan Toronto Social Planning Council (1984) Caring for Profit: the commercialization of human services in Ontario. Toronto, Ontario.

Science 3 February 1984, p. 463.

Shannon G., Dever, G. (1974) Health Care Delivery: spatial perspectives New York: McGraw Hill.

Streiner D.L. (1985). Teaching Critical Appraisal Skills (in this volume).

Time, 10 December 1984, pp. 42-69.

Tsai L., Tsuang M.T. (1981). How can we avoid unnecessary CT scanning for psychiatric patients? Journal of Clinical Psychiatry 42, 452-454.

TEACHING CRITICAL APPRAISAL SKILLS

DAVID L. STREINER Ph.D.

ABSTRACT

The ability to appraise scientific articles is an indispensable skill. We are long past the day when what was learned in school would stand us in good stead throughout our careers. This paper presents in detail, a course from McMaster University on the critical appraisal of "Scientific" literature. The content of the nine two-hour sessions is described.

By now, it has become almost commonplace to talk about our inability to keep up with the amount of information that appears in scientific journals. It has been estimated that the quantity of all learned literature shows a growth of about 6 or 7% a year, doubling in size every 10 to 15 years (de Solla Price, 1975). About 70 years ago, there were approximately 1,000 biomedical journals: 50 years ago, 1,500 serials; 4,000 30 years ago; 10 years ago there were 14,000; and now there are well over 20,000 biomedical journals published annually (Warren, 1981). To take just one example of how this affects the practising physician, Riley (1984) estimated that in the 10 years between the fourth and sixth editions of The Pharmacological Basis for Therapeutics, nearly 25% of the generic drug names were dropped, and one-third of the 1980 list was not present in 1970. The clinician faces a formidable challenge: how to keep abreast of current knowledge, while the amount of that information is growing exponentially. Further complicating the task is that the quality of articles varies tremendously, ranging from understated ones that have major implications for clinical practice, to reports of highly-publicized trials which turn out to not meet minimal standards of scientific rigour.

The busy clinician needs to be able to read journals quickly and efficiently, in order to "separate the wheat from the chaff" (Streiner, 1983); that is, to learn how to rapidly spot articles which do not meet certain methodological criteria and can be discarded; and to recognize those which do, and thus should influence his or her clinical practice. The approach taken by the Department of Clinical Epidemiology and Biostatistics at McMaster University has been to offer a brief, two hour per week course over nine weeks, whose aims are presented in Table 1. Let me emphasize what I believe are the key elements: the major focus is on the application of these skills in clinical practice rather than on teaching statistical principles; it attempts to make the use of these tools an integral part of practice, rather than another corpus of arcane but rarely applied knowledge; and it is designed as much for the "consumer" of research as for the "doer".

TABLE 1

Objectives of the Critical Appraisal Course

1. To allow participants to acquire an efficient, critical approach to making clinical decisions based on an objective assessment of the literature.

2. To demonstrate how this approach is complementary to bedside teaching, and can be incorporated into the care of individual patients.

3. To provide exposure to research concepts and show how a knowledge of methodological issues is important for "consumers" of research as well as "doers" of research.

4. To identify participants who might enjoy doing research.

During this talk, I would like to cover three topics: the content of the courses; how a typical course is put together; and our experience with them to date. First, though, a bit of historical background. The Critical Appraisal of Clinical Evidence course was first given by Drs. David Sackett and Brian Haynes as a series of lunchtime, "eat and learn" sessions in 1978, with an audience made up primarily of clinical faculty and community physicians. Since that time, specialty-specific courses have been developed for Family Medicine, Obstetrics and Gynecology, Internal Medicine, Pediatrics, Psychiatry, Radiology, Rheumatology, Clinical Pathology, Surgery, Nursing, and Physiotherapy.

The Content of the Course

As I mentioned earlier, a typical course is nine weeks long, during which time the topics listed in Table 2 are covered. Each topic is built around one of the articles written by various members of the Department and published in the Canadian Medical Association Journal under the main headings of "How to Read Clinical Journals," "Clinical Disagreement," and "Interpretation of Diagnostic Tests" (see Appendix A for a full list of titles). The students also have two or three articles which address the issue under discussion, and some worksheets to help them evaluate the articles.

TABLE 2

Format of the Critical Appraisal Course

Session	Topic
1 and 2	How to assess studies that evaluate the effectiveness of a specific therapy.
3	How to assess studies that evaluate diagnostic tests.
4	How to assess studies that evaluate clinical error.
5	How to assess studies that evaluate a patient's prognosis.
6	How to assess studies that evaluate etiology and risk factors.
7	How to assess studies that evaluate the quality of medical care.
8	How to assess efficiency studies; i.e., studies that evaluate cost-effectiveness, cost-benefit, and cost-utility of medical care.
9	Students present topic of their choice.

The goal of sessions 1 and 2 is to enable the students to determine if an article advocating a specific treatment has drawn conclusions about efficacy that are both valid and applicable in their clinical settings. In reading the articles, the students are asked to evaluate the studies according to the criteria outlined in Table 3. I'll go over these in some detail so you can get a flavour of how we approach each session. The first criterion, random assignment to treatment groups, is perhaps the most critical standard we insist upon. Although we recognize that the Randomized Control Trial (RCT) is not the only golden road to truth, it is perhaps the most valid. Studies which do not rely on random assignment are susceptible to a host of biases which make the results questionable at best, and totally invalid at worst. The use of historical control groups or reports of a series of cases from one centre cannot rule out the effects of changes in clinical practice over time, placebo or Hawthorne effects, or other factors which may account for the reported results. Our favourite examples are gastric freezing for peptic ulcers and mammary ligation for angina; thousands of these procedures were performed on hapless patients, based on ecstatic reports of uncontrolled trials, until RCTs demonstrated the ineffectiveness of these quite radical techniques (Miao, 1977; Cobb, Thomas, Dillard, Merendino, & Bruce, 1959).

Table 3

Criteria for Evaluating Effectiveness of Therapy

1. Was the assignment of patients to treatments really randomized?
2. Were all clinically relevant outcomes reported?
3. Were the study patients recognizably similar to your own?
4. Were both statistical and clinical significance considered?
5. Were all patients who entered the study accounted for at its conclusion?
6. Is the therapeutic maneuver feasible in your setting?

For criterion two, we ask the students to put down the article before they reach the conclusions section and to ask themselves; as clinicians, what outcomes would they deem important if they were evaluating this new therapy? After they have made up their list, compare it to what the investigators looked at -- did the researchers play a "substitution game," looking at, for example, a reduction in diastolic BP in a group of Type A executives, when they were really interested in a reduction in the number of fatal or non-fatal MIs or CVAs? Similarly, did they look at outcomes that patients would consider important? Our pet example of these two points is the clofibrate trial (Oliver, Heady, Morris, & Cooper, 1978). The authors reported significant decreases in serum cholesterol and the number of fatal and non-fatal MIs in men treated with this drug as compared with a placebo group. From the point of view of the researchers, who were interested in the effects of clofibrate on re-infarctions, these results were extremely promising. However, not all patients would have agreed, as the total number of deaths from all causes was significantly higher among clofibrate users.

The third criterion asks whether the patients in the study are similar to those whom the readers see. For example, the results from a trial which took place in a tertiary care setting, with subjects who were carefully screened to eliminate the potential non-compliers, may have little relevance in a family practice setting; and vice versa.

Criterion four differentiates between the statistical significance and the clinical importance of the results. Schwartz and Shapiro (1973), for instance, found a statistically significant reduction in diastolic blood pressure as a result of a biofeedback program. However, the magnitude of the decrease was only 5mm Hg, which is clinically useless for clinicians treating primary hypertensives.

Criterion five asks whether all of the patients who began the trial were accounted for at the end. Naturally, some subjects move, others want no more to do with the experiment, and some simply drop out of sight. If this happens differentially in one group, or occurs in more than 15% of the

cases, we would question the validity of the results. A conservative check would be to assume that all drop-outs had poor outcomes; if the therapy still looks good, then fine; but if the new procedure now yields equivocal or negative results, we would ungraciously consign the article to the circular file.

Even if the study passes the first five criteria, it may still be infeasible in the reader's setting. The equipment may not be available; the procedure may be too costly; or the patients may find it too expensive, invasive, or discomforting. Thus, a therapy must not only be efficacious, it must be practical.

The third session looks at articles which propose new diagnostic tests, although our aim is to also have the students re-evaluate the tests they now use to determine their utility and accuracy in the students' own clinical settings. Here, we propose the eight criteria listed in Table 4. Without going through each of the criteria individually, our aim is two-fold: to have the students evaluate how the test was developed; and to see if it will be accurate with their mix of patients.

TABLE 4

Criteria for the Evaluation of a Diagnostic Test

1. Was there an independent, blind comparison with a "gold standard" of diagnosis?

2. Did the patient sample include an appropriate spectrum of mild and severe, treated and untreated disease, plus individuals with different but commonly confused disorders?

3. Was the setting for the study, as well as the filter through which the study patients passed, adequately described?

4. Was the reproducibility of the test result (precision, reliability) and its interpretation (validity) determined?

5. Was the term "normal" described sensibly?

6. If the test is advocated as part of a cluster or sequence of tests, was its contribution to the overall validity of the cluster or sequence determined?

7. Were the tactics for carrying out the test described in sufficient detail to permit their exact replication?

8. Was the utility of the test determined?

For example, was the test standardized on a randomly selected group of normals and patients, or did the author use himself and 12 colleagues

and technicians who happened to be around the lab on a particular day? Were the patients typical of those for whom the test is proposed, or were they a highly selected group of people in a tertiary care setting, whereas the procedure is proposed for use as a screening test in a primary care setting? The bulk of the session, though, is devoted to the issue of the effects of the prevalence of the disorder on the test's utility. In a classic paper that is now 30 years old, Meehl and Rosen (1955) pointed out that a test may look extremely promising when a group of patients with the disorder is compared to another group of patients of equal size who do not have the disorder. However, if the prevalence of the disorder is significantly different from 50%, then the test will yield an unacceptably high level of false negative or false positive findings.

Let me illustrate this briefly. Carroll, Curtis, and Mendels (1976) proposed the dexamethazone suppression test (DST) for diagnosing endogenous depression, using the data summarized in Table 5 to support their claim that it could discriminate between these and other psychiatric patients. Indeed, according to their results, the positive predictive value of the test is 83% -- that is, of all positive lab results, 83% will correctly identify depressed patients; only 17% will erroneously label a non-depressed person as depressed (although 43% of depressed patients will incorrectly be called non-depressed).

TABLE 5
Utility of DST, From Original Article

		Criterion Diagnosis		
		Endogenous Depressives	Other Disorders	
DST Result	Positive	24	5	29
	Negative	18	37	55
		42	42	84

Prevalence = 50% Positive Predictive Value = 24/29 = 83%

Adapted from Carroll et al., 1976.

However, if the test is used on all new admissions to a psychiatry unit where the prevalence of endogenous depression is only 10% (as was being done in some hospitals), then the results look quite different. As

Table 6 shows, now only 35% of the positive test results are from depressed patients; the other 65% are all false positives; and we would still miss 43% of the true cases. (Just so our psychiatric colleagues wouldn't feel maligned or inadequate, we have shown similar results with tests from every specialty, ranging from SMA-12 to stress EKGs.)

Table 6

DST Results with 10% Prevalence of Depression

		Criterion Diagnosis		
		Endogenous Depressives	Other Disorders	
DST Result	Positive	57	107	164
	Negative	43	793	836
		100	900	1000

Prevalence = 10% Positive Predictive Value = 57/164 = 35%

The third set of readings looks at clinical disagreement, or why two clinicians can come to different conclusions regarding the presence or absence of a clinical sign or symptom. Here we address variations arising from any of three sources: the examiner, the examined, and the examination (see Table 7). In addition to covering how variability may arise in each of these areas, we also look at how agreement among examiners is measured, and how we can get spuriously high levels of concordance if we ignore the effects of agreement by chance. To take a simple example, let's assume that two clinicians are asked to independently state whether each of 100 patients is schizophrenic or not. Both of them, though, feel overwhelmed by other work and do not have time for this study, so they say to themselves, "Of the patients I've seen here, roughly 30% have been schizophrenic, so I'll just randomly call 30% schizophrenic and 70% non-schizophrenic, without even bothering to look at them." Adopting this strategy, they will have an overall agreement rate of 58%--they will come up with the same diagnosis for nine of the schizophrenics and 49 of the 70 non-schizophrenics. Thus, any real study which found 65% agreement between two clinicians in a similar setting--not an unusual level in psychiatry--must be looked at in light of the 58% agreement by chance alone.

TABLE 7
Sources of Measurement Error

The Examiner

- Biological variation in the senses
- Recording inferences, not evidence
- Ensnarement by diagnostic classification schemes
- Entrapment by prior expectations

The Examined

- Biological variation in the observed system
- Effects of illness and medication
- Memory and rumination

The Examination

- Disruptive environments
- Disruptive interactions
- Incorrect use or function of diagnostic tools

The fourth series of readings is devoted to reading articles regarding the prognosis of a disorder. The guidelines the course proposes are given in Table 8. The primary emphasis in this session is on how the sample was gathered originally, and how complete the follow-up was. Unless patients were gathered at an early and uniform time in the course of their disorder (an "inception cohort"), data regarding prognosis (and prevalence, for that matter) are highly suspect. For example, if we were interested in the prognosis of people with psychotic depressions, one common strategy would be to interview all patients who are admitted to an inpatient unit with this diagnosis, regarding previous depressive episodes they may have had. Since patients who have numerous and frequent episodes have many more chances to be in hospital at the time of the study, as compared with people who have few or no recurrences, our results may look quite bleak. The reasons for this are obvious -- because of our sampling strategy, patients with good prognoses are under-represented in our sample.

TABLE 8

Criteria for Evaluating Clinical Course and Prognosis

1. Was an inception cohort gathered?
2. Was the referral pattern described?
3. Was complete follow-up achieved?
4. Were objective outcome criteria used?
5. Was the outcome assessment blind?
6. Was there adjustment for extraneous prognostic factors?

On the other hand, our data may be unrealistically optimistic, as there is another group which is not sampled at all. These are the patients who, during or following a depressive episode, commit suicide. Thus, our sample would not contain any patients whose depression was so severe as to be fatal.

Where between these two prognoses--one artifactually poor and the other artificially good--does the "real" prognosis lie? The answer is simply that we don't know, and can't know from this type of study, as we did not begin with an inception cohort.

The next set of readings deals with articles which attempt to determine etiology or causation; for example, that schizophrenia is determined by genetic as opposed to environmental factors, that depression is caused by a disruption in the circadian rhythm, or that tardive dyskinesia is a result of the chronic use of phenothiazines. Here we apply the criteria outlined in Table 9. Criterion 1 focusses on the strength of the experimental design; was it an RCT, a cohort study, a case-control design, or simply a series of case reports? Needless to say, the stronger the design, the more believable are the conclusions regarding causality. The remaining criteria look at the strength, consistency, and sensibility of the association between the putative cause and the disorder. In brief, they ask whether the results are consistent with previous studies, our knowledge of the epidemiology and biology of the disease, and if some of the basic laws of logic apply--cause should precede result and, in most cases, more of the cause should lead to more of the result.

TABLE 9

Criteria for Evaluating Causation

1. Is there evidence from true experiments with humans?
2. Is the association strong?
3. Is the association consistent from study to study?
4. Is the temporal relationship correct?
5. Is there a dose-response gradient?
6. Does the association make epidemiological sense?
7. Does the association make biologic sense?
8. Is the association specific?
9. Is the association analogous to a previously proven causal association?

Lest we think that these criteria apply only to the "hard science" aspects of medicine, let me cite a few examples from psychology and psychiatry. With respect to the dose-response criterion, we would expect that, if there is a genetic component to schizophrenia, then kids with one schizophrenic parent should be at higher risk than kids with no schizophrenic parent, and having both parents so diagnosed should increase the risk even more. This in fact was what was found by Gottesman (1979). Similarly, concordance rates among siblings should be higher for monozygotic twins than dizygotic twins; again, this is consistent with the literature (Gottesman & Shields, 1972).

On the other hand, the theory that early infantile autism is caused by the emotional frigidity of the parents (e.g., Eisenberg & Kenner, 1956) derives from case-control studies, where the children were identified by current symptomatology, and "causes" are derived from history-taking. Thus, cause and effect are measured simultaneously, with the direction of causality imposed by theory, not the data. With the same data, Peck, Rabinovitch, and Cramer (1949) postulated that the emotional distancing of the parents was a reaction to their infant's unresponsiveness, not the cause. So, the need for well-designed studies to determine etiology exists as much in psychiatry as in other branches of medicine.

The major focus of the sixth topic is on articles which attempt to measure quality of care through chart audits, using the criteria outlined in Table 10. By the end, the student should be aware of what cannot be assessed through chart review as well as what can be. For example, we are all aware that an entry in a chart usually means that that procedure was done. On the other hand, the absence of an entry does not necessarily mean that the procedure was not done -- it may have been performed but not charted. Similarly, the chart entries usually end when the patient is discharged. Consequently, if a procedure results in a disastrous, but delayed, outcome, or if it causes catastrophic outcomes in only a small proportion of the cases, an audit may not pick it up. So, chart reviews are pretty useless if they are evaluating only the outcome of an intervention.

TABLE 10
Criteria for Evaluating Quality of Care

1. Did the study focus on what clinicians actually do?

2. Have the clinical actions been shown to do more good than harm?

   2a) If not, is the article about a process-versus-outcome study that assembled an inception cohort, ensured high rates of compliance, included all relevant outcomes, and considered prognosis?

3. Are the clinicians, patients, and type of practice similar to your own?

4. Were the clinical actions measured in a clinically sensible fashion?

5. Were the clinical actions measured in a scientifically credible fashion?

6. Were both clinical and statistical significance considered?

On the other hand, well conducted audits may be quite good if they are aimed at the processes of care. Once we mention "well conducted," we are back into the realm of the design of the study, and we translate criteria used in previous sessions to chart review. Let me emphasize that chart audit can be more than a tedious procedure done merely to satisfy the accreditation team; if it is well-designed and well-executed, it can yield potentially useful information.

The last set of readings is concerned with economic evaluations. Once it has been shown that an intervention can work under ideal conditions (efficacy), does work in actual practice (effectiveness), and is reaching those who need it (availability), clinicians -- and often administrators -- begin to ask questions like, "How much is this costing," or "Can I get equivalent results for less money" (McNamara's "more bang for the buck" philosophy, much beloved by the Department of Defence), or "Into which of two competing, and very different programs, should we invest our limited money." Since this is often terra incognita for clinicians, much of the session is devoted to simply introducing them to the various terms and concepts, such as those presented in Table 11.

TABLE 11

Terms Used in Health Care Evaluation

<table>
<tr><td colspan="2"></td><td colspan="3">Are both costs and consequences of the alternatives examined?</td></tr>
<tr><td rowspan="3">Comparison of two or more alternatives?</td><td rowspan="3">NO</td><td colspan="2">No</td><td>Yes</td></tr>
<tr><td>Examines only Consequences</td><td>Examines only Costs</td><td></td></tr>
<tr><td>Partial Evaluation
Outcome Description</td><td>Partial Evaluation
Cost Description</td><td>Partial Evaluation
Cost-Outcome Description</td></tr>
<tr><td>YES</td><td colspan="2">Partial Evaluation
Efficacy or Effectiveness Evaluation
Cost Analysis</td><td>Full Evaluation
Cost-minimization
Cost-effectiveness
Cost-benefit
Cost-utility
Analysis</td></tr>
</table>

The other key point emphasized in this session is the point of view of the evaluation. That is, a program which transfers psychiatric in-patients to day-hospital status may look very good from the hospital's point of view, in terms of generating more revenue with perhaps fewer staff. However, patients and their families may experience an economic loss, due to transportation costs, lost income for family members who have to drive the patient to and from the hospital, and having to pay for medications they have been receiving free of charge while in hospital.

The final session has two main goals. First, one or two students are encouraged to bring an article they feel should be evaluated, preferably handed out at the end of the previous class. They then lead the discussion, so they can try out their newly gained skills on their own. In some sessions, students have brought articles they themselves have written or, better yet, grants which they plan to submit. The second goal is less didactic, but perhaps more educational: to split some pizzas and a bottle of wine.

How to Put a Course Together

It is important to remember that the outline that I've presented is not written in stone; the contents can be varied to suit the intended audience. For example, in some courses I've given for psychiatric residents, I've added a session in the beginning covering some of the basic concepts of statistics. These people have either never had any introduction to statistics, or have forgotten what they had been taught and felt that they needed just an introduction to the various tests they would be encountering in the articles.

Once the topics are chosen, we usually bring a committee of about seven or eight people together, one person to cover each of the sessions. Their job is outlined in Table 12.

TABLE 12
Designing a Course

1. Determine the objectives for each session.

2. Choose two or three articles that are typical in that area.

   2a) Be prepared to replace the articles when you or your students come across better ones.

3. Write a scenario, presenting the student with a clinical problem which has to be answered using the articles.

4. Prepare a tutor's guide, answering the questions in the worksheet.

The first task is to define the objectives of that session. These can be taken from the ones that we have derived, or modified to suit your setting and audience. To make the session more meaningful, we introduce it with a clinical scenario, such as in Table 13, which we use for the session on Therapy.

TABLE 13
Scenario for Therapy Session

You see a patient, Mrs. Heartstone, who has a primary diagnosis of depressive psychosis. She is currently on imipramine, 50 mg. t.i.d., but she is not doing well, exhibiting an extremely depressed mood, suspiciousness, and ideas of bodily change when you talk to her. You wonder whether you should refer her for ECT. She has not received any ECT since she was first diagnosed six years ago, and is physically well otherwise.

Read the attached reprints:

Freeman, C.P.L., Basson, J.V. (1978). Double blind controlled trial of electro-convulsive therapy (ECT) and simulated ECT in depressive illness. The Lancet, 8 April, 738-740.

Johnstone, E.C. et al. (1980). The Northwick Park electro-convulsive therapy trial. The Lancet, 20 December, 1317-1320.

"How to read a clinical journal. V: To distinguish useful from useless or even harmful therapy."

This presents a problem to the students, typical of a clinical situation they may actually encounter in their daily practice. In order to answer the question, the students are provided with two or three articles from the literature which address the issue. These articles are not chosen because they are particularly good or bad; we try to choose ones that are representative of those the students will encounter when they open a journal they are accustomed to reading. This is where it is important to select your committee members carefully -- people who know the literature in their area and are able to put their fingers on some typical articles. Bear in mind that on your first go-round, the articles you select won't be ideal; they may turn out to be too turgid, have some glaring faux pas that may distract the students from the points you want to get across, and so on. Be prepared to replace the articles; again using the course we put together in Psychiatry, only half of the articles we now use were the ones we originally selected five years ago. Further, the articles I use when I teach the course in another setting are slightly different from the ones I use at home; because of their particular expertise, they are more interested in articles in one area--anorexia nervosa and bulimia--than are residents in another setting.

The third task is to select one or two methodological articles which can help them evaluate the clinical readings. Needless to say, we use the ones written by various members of our Department. However, there are other guides available, which could be chosen if they better meet your needs.

The last, and perhaps hardest, job is to provide a tutor's guide. We don't expect that the original committee will be the only ones teaching the course; if your experience is anything like ours, the demand will grow, and you will need new people. We have also found it a good idea to involve new people periodically, simply because they bring in a fresh perspective. As I mentioned before, the students are given a worksheet for each article, posing the questions outlined in the criteria for that session. The tutor's guide is a "gyp sheet," giving the answers to the questions, and the rationale for each answer. In this way, we can use the old medical school dictum of "see one, do one, teach one." All of the tutors we use now have been former students in the course, who then went on to be co-tutors with more experienced people, and then took over the course on their own. The tutor's guide makes this transition easier, and ensures some degree of continuity and consistency from one tutor to the next. At the end of each session, the guides are given to the students, to which they can refer later to see our mistakes.

Our Experience with the Course

As I've already mentioned, the Critical Appraisal course has grown from an "undifferentiated," lunchtime course to a number of specific ones covering most specialty areas in medicine and other health-related disciplines. For the past four years, we have held one-week workshops for people from other institutions, so that they could learn how to prepare these courses in their settings. Other workshops have been held at the Society for Research in Primary Care in Internal Medicine; in Brazilia; and at the Sechuan Medical College. As of June of 1983, over 2,000 faculty, residents, undergraduates, nurses, physiotherapy students, and visitors have taken the course. Contact with former students at other settings is kept alive through the International Resource Bank; they can write to see what articles are used here and elsewhere, and are encouraged to send in their versions of the course.

It would be somewhat hypocritical of us to advocate that the use of any new intervention should be predicated on its proven effectiveness, and not to evaluate the utility of these courses. In 1980, Kathy Bennett and some other members of the Department did just this, using clinical clerks in our program (Bennett, Haynes, Neufeld, Tugwell, & Sackett, 1984). Tutors at two of the teaching hospitals were given special training sessions in the use of the critical appraisal packages; the tutors in the other two hospitals served as the control condition. All of the clerks assigned to these tutors did a paper-based evaluation at the beginning and end of the rotation. These problems were similar to ones found in the Therapy and Diagnosis critical appraisal sessions. Briefly, the results were that, not only did the "experimental clerks" increase in their ability to critically appraise therapy and diagnosis problems, the clerks in the control condition actually decreased. Bennett is now undertaking a more elaborate study of critical appraisal, growing out of her Master's thesis. Hopefully, the results of this should be known soon.

In conclusion, we feel that the ability to appraise what is in the literature is an indispensable skill that all health-care practitioners should possess. We are long past the day when what was learned in school would stand us in good stead throughout our careers. To keep up with current practice, we must read; but read in a critical way so that we do not blindly accept every new fad or unproven therapy. This is especially true when new diagnostic tests and interventions are becoming increasingly expensive, and yet hold out the promise of much improved care. New advances in biotechnology will best succeed if they are coupled with improvement in our ability to assess them.

REFERENCES

Bennett, K.B., Haynes, R.B., Neufeld, V.R., Tugwell, P., & Sackett, D.L. (1984). A controlled trial of teaching critical appraisal of the medical literature to M.D. clinical clerks. McMaster University.

Carroll, B.J., Curtis, G.C., & Mendels, J. (1976). Neuroendocrine regulation in depression: II. Discrimination of depressed from nondepressed patients. Archives of General Psychiatry, 33, 1051-1058.

Cobb, I.A., Thomas, G.I., Dillard, D.H., Merendino, K.A., & Bruce, R.A. (1959) An evaluation of internal-mammary artery ligation by a double-blind technic. New England Journal of Medicine, 260, 1115-1118.

de Solla Price, D. (1975). Science Since Babylon. New Haven: Yale University Press.

Gottesman, I.I. (1979). Schizophrenia and genetics: Toward understanding uncertainty. Psychiatric Annals, 9, 26-37.

Gottesman, I.I., & Shields, J. (1972). Schizophrenia and genetics: A twin study vantage point. New York: Academic Press.

Meehl, P.E., & Rosen, A. (1955). Antecedent probability and the efficacy of psychometric signs, patterns, or cutting scores. Psychological Bulletin, 52, 194-216.

Miao, L.L. (1977). Gastric freezing: An example of the evaluation of medical therapy by randomized clinical trials. In J.P. Bunker, B.A. Barnes, & F. Mosteller (Eds.), Costs, risks, and benefits of therapy (pp. 198-211). New York: Oxford University Press.

Oliver, M.F., Heady, J.A., Morris, J.N., & Cooper, J. (1978). A co-operative trial in the prevention of ischaemic heart disease using clofibrate: Report from the Committee of Principal Investigators. British Heart Journal, 40, 1069-1118.

Riley, M.W. (1984). Reducing "information overload" in the teaching of pharmacology: The "200 drug list." Journal of Medical Education, 59, 508-511.

Schwartz, G.E., & Shapiro, D. (1973). Biofeedback and essential hypertension: Current findings and theoretical concerns. In L. Birk (Ed.), Biofeedback: Behavioral medicine, (pp. 133-143). New York: Grune and Stratton.

Streiner, D.L. (1983). Separating the wheat from the chaff: How to locate and use the best available literature. In D. P. Lumsden (Ed.), Community mental health action: Primary prevention programming in Canada (pp.62-69). Ottawa: Canadian Public Health Association.

Warren, K.S. (1981). Selecting aspects of the biomedical literature. In K.S. Warren (Ed.), Coping with the biomedical literature: A primer for the scientist and the clinician (pp. 17-32). New York: Praeger.

APPENDIX A

ARTICLES IN THE CRITICAL APPRAISAL SERIES

1. Clinical Disagreement Series

Department of Clinical Epidemiology and Biostatistics. (1980). Clinical disagreement. I: How often it occurs and why. Canadian Medical Association Journal, 123, 499-504.

Department of Clinical Epidemiology and Biostatistics. (1980). Clinical disagreement. II. How to avoid it and how to learn from one's mistakes. Canadian Medical Association Journal, 123, 613-617.

2. How to Read Clinical Journals Series

Department of Clinical Epidemiology and Biostatistics. (1981). How to read clinical journals. I: Why read them and how to start reading them critically. Canadian Medical Association Journal, 124, 555-558.

Department of Clinical Epidemiology and Biostatistics. (1981). How to read clinical journals. II: To learn about a diagnostic test. Canadian Medical Association Journal, 124, 703-710.

Department of Clinical Epidemiology and Biostatistics. (1981). How to read clinical journals. III: To learn about the clinical course and prognosis of disease. Canadian Medical Association Journal, 124, 869-872.

Department of Clinical Epidemiology and Biostatistics. (1981). How to read clinical journals. IV: To determine etiology or causation. Canadian Medical Association Journal, 124, 985-990.

Department of Clinical Epidemiology and Biostatistics. (1981). How to read clinical journals. V: To distinguish useful from useless or even harmful therapy. Canadian Medical Association Journal, 124, 1156-1162.

Department of Clinical Epidemiology and Biostatistics. (1984). How to read clinical Journals. VI: To learn about the quality of clinical care. Canadian Medical Association Journal, 130, 337-381.

Department of Clinical Epidemiology and Biostatistics. (1984). How to read clinical journals. VII: To understand an economic evaluation (Part A). Canadian Medical Association Journal, 130, 1428-1433.

Department of Clinical Epidemiology and Biostatistics. (1984). How to read clinical journals. VIII: To understand an economic evaluation (Part B). Canadian Medical Association Journal, 130, 1542-1549.

3. Interpretation of Diagnostic Data Series

Department of Clinical Epidemiology and Biostatistics (1983). Interpretation of diagnostic data. I: How to do it with pictures. Canadian Medical Association Journal, 129, 429-432.

Department of Clinical Epidemiology and Biostatistics (1983). Interpretation of diagnostic data. II: How to do it with a simple table (Part A). Canadian Medical Association Journal, 129, 559-564.

Department of Clinical Epidemiology and Biostatistics (1983). Interpretation of diagnostic data. III: How to do it with a simple table (Part B). Canadian Medical Association Journal, 129, 705-710.

Department of Clinical Epidemiology and Biostatistics (1983). Interpretation of diagnostic data. IV: How to do it with a more complex table. Canadian Medical Association Journal, 129, 832-837.

Department of Clinical Epidemiology and Biostatistics (1983). Interpretation of diagnostic data. V: How to do it with simple maths. Canadian Medical Association Journal, 129, 947-954.

Department of Clinical Epidemiology and Biostatistics (1983). Interpretation of diagnostic data. VI: How to do it with more complex maths. Canadian Medical Association Journal, 129, 1093-1099.

DISCUSSION and SYNTHESIS

J.M. CLEGHORN, M.D.

Editor's note

Dr. J.M. Cleghorn had been asked to participate in the symposium and prepare some Syntheses, adding his own perspective.

He spoke on two occasions. The first followed Part II and is on page 189 of this document, entitled Brain-Imaging with PET-A Scientific Bridge.

In these final remarks, concluding the Symposium, he refers not only to material published here but also to contributions made by Dr. Patrick McGeer and Dr. Stuart Smith. The editors regret that their very thoughtful contributions were not available for publication. Nevertheless Dr. Cleghorn includes some highlights from their presentations.

Dr. Cleghorn

First, I want to thank the organizers of the conference for the privilege of listening to an excellent set of presentations. This has been a first class educational experience and the publication which results will certainly meet international standards for excellence and will provide critical update of major issues in clinical psychiatry, biotechnology and related socioeconomic issues.

There are certain points of practical information that should be considered for wide dissemination to practising mental health workers: (1) familial information on alcoholism; (2) key questions that practitioners can employ in order to detect alcoholism and the two lab tests mentioned; (3) the new treatments for anxiety disorders and their efficacy. This information has been made very available to psychiatrists and it would make an interesting project to determine how much public awareness there is of this information and how much awareness amongst family practitioners and other mental health workers and, if needed, an appropriate educational program mounted.

It should be a matter of pride that the presentations were excellent and meet international standards. This is not simply a matter of national pride, which Dr. Dear mentioned this morning, as possibly a faulty source of motivation for achievement. Rather, I believe that the appeal by Doctors Mustard and Smith that we should achieve our potential in science in this country is actually a thinly disguised way of talking about shared personal pride and, in my view, this is legitimate pride because it is concerned with achieving one's best and being given the opportunity to make one's best contribution.

The pace of progress in technology is so rapid that it seems virtually certain there will be opportunity to do research which may change our perspectives on major mental illnesses radically in the next few years. Whether this will offer safe and practical treatments is very hard to say but the pace of progress is such that it would be a great pity to delay it now. This country is not by any means weak in neurobiology. Huge investments have been made over the last several decades and many who have gone before us have dedicated their lives to preparing the field for the opportunities we have now.

If there are major gaps in Canadian neurosciences relevant to mental disorder, they were not identified at this meeting.

There are, however, some problems. For example, the insecurity of funding is a serious problem to some young investigators, discouraging them from pursuing the field. In addition, clinical investigation in psychiatry in some centres will need to be strengthened if we are to succeed. In a survey I made of psychiatry departments over a year ago, it appeared that positions were available for such people. One or two in each centre is all that would be required at present. A few students in our medical schools are beginning to perceive these opportunities and we are beginning to attract some of the brightest into research in neurosciences related to psychiatry. It will take a number of years before these people can join us and there is an opportunity before us now that must be grasped. In the United States, the same trends have occurred, at Harvard, for example. There are increasing numbers of excellent students trying to make their way into the field. It is difficult for them in view of cutbacks in research funding in the United States which are taking place at the present time. Why shouldn't we make an opportunity for a few of these people to express their talents and, at the same time, to strengthen ourselves. I do not believe it would take a massive infusion of funds to increase our forward thrust in this way, to mount a national program against the enemy in inner space.

Dr. Dear argued that excessive emphasis on technology ignores some very practical social health problems, such as housing. I think it is important to distinguish between excessive expenditures on technology in the health care system (e.g. C.A.T. scanners in every village in the nation) and technology to produce new knowledge which may result in inexpensive treatments. In addition, technology in general medicine has generated staggering costs by lengthening the lives of the disabled. In psychiatric illness this is unlikely to happen. Our patients are disabled but most of them do not die young. Their disability generates huge costs and it will be important to assess whether in fact new discoveries in the next two decades lower the cost of mental illness.

Dr. McGeer added a new twist to the position that health care must be accessible to all citizens in Canada. He said that those suffering from neuropsychiatric illnesses can't make a legitimate claim on the health care system because there are no effective treatments. Taken to its extreme, this argument would suggest that it is unconstitutional not to give support to research to solve these problems. Perhaps Dr. McGeer was saying, in

more modest terms, that we are so near and yet so far from important discoveries. He emphasized that imaging, though very expensive, is the only way to answer certain questions. Dr. Mustard thought we might not make it on the world scene. Lacking a national policy, we have diffused our resources with the excellent technical resources at UBC and an outstanding investigative resource at McMaster with a less adequate imaging facility. He called this a classic example of misallocation of resources and attributed it to the lack of a national policy. He called for the need to concentrate resources in certain centres. Alternatively, a consortium could be developed such as he has developed under the aegis of the Canadian Institute for Advanced Research. Sensible and innovative as these ideas are, they are almost un-Canadian. We have no mechanism for establishing a national policy or for concentrating resources. In Ontario we have more universities than funds available to maintain their quality, yet it is unthinkable that one of them might be terminated in order to save the quality of the others. In Canada we place a higher value on fairness, peace and regional equality than on achievement.

Of the many interesting points made by Dr. Stuart Smith, one was the need for a pharmaceutical industry and a medical devices industry in Canada. This provoked some misunderstanding that perhaps science would then be driven by industrial requirements. Dr. Smith's point, I believe, is that when science produces a marketable product we should have the capability of producing it and selling it and taking the profits. The reason for this is that the country needs to generate more wealth in order to support its science, in order to make the best use of the talent in the scientific community. Furthermore, such possibilities of profit constitute an incentive to potential private donors. There are many potential donors who are prepared to commit themselves to long term research and development, providing they can see some long term possibility of financial self maintenance.

Mr. Moriarity criticized Dr. Smith for not having come to terms with the realities of living in Canada. He stated that we did have a pharmaceutical industry but that we do not own it. He cited instances of numerous university professors who have set up assay systems that have become small businesses. He did comment, however, that when the opportunity to market one of the products in the United States became available, this was impossible because of the lack of financial investment required in order to develop a market system. Rather, the small company in question had to enter into a contractual agreement with a distributor in the United States. I think this view represents an acceptance of Canadian realities which some of us have trouble accepting.

If we do not manage to influence our government to make a concerted thrust in the high technology of applied neurosciences, then we must select small feasible and intelligently conceived projects. Many were reported on in this meeting and many will carry on unaffected by the outcome of the debate recorded here. I suspect that may be the realistic choice, but with regret that we may not be able to make the full contribution of which we are capable.